Tuomas Leinonen

Aligning and tracking a beam steerable millimeter-wave radio link

School of Electrical Engineering

Thesis submitted for examination for the degree of
Master of Science in Technology.

Espoo, Finland September 23rd, 2015

Thesis supervisor:

Prof. Ville Viikari

Thesis advisor:

Jyri Putkonen, M. Sc. (Tech.)

Aalto University
School of Electrical
Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Tuomas Leinonen

Title: Aligning and tracking a beam steerable millimeter-wave radio link

Date: September 23rd, 2015 Language: English Number of pages: 15+136

Department of Radio Science and Technology

Professorship: Radio Science and Engineering

Supervisor: Prof. Ville Viikari

Advisor: Jyri Putkonen, M. Sc. (Tech.)

In order to provide high-throughput mobile broadband in a dense urban information society, upcoming cellular networks will finally employ the under-utilized millimeter-wave (mmW) frequencies. The challenging mmW radio environment, however, necessitates massive cell densification with wireless backhauling using very directional links. This thesis investigates how these links between access points may be aligned efficiently, and how alignment reflects the network organization.

The work provides a thorough presentation of different high-level aspects and background information required when designing a mmW small cell system. In terms of alignment functionality, both automatic link establishment and proactive tracking are considered. Additionally, the presentation includes an overview of beam steerable antennas, mmW propagation in urban environments, and network organization. The thesis further specifies requirements, proposes possible approaches and compares those with existing implementations.

Most of existing mmW beam alignment solutions are intended for short-range indoor communications and do not address the issues in cellular systems. While existing functionality considers only a single link between two devices, efficient design should consider both the entire network and the underlying phenomena. The devices should further exploit the existing network infrastructure, location and orientation information, and the concepts of machine learning. Even though the world has recently seen advancements in the related fields, there is still much work to be done before commercial deployment is possible.

Keywords: 5G, algorithm, antenna, beam alignment, beamforming, beam steering, beam tracking, millimeter wave, radio, radio wave propagation

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Tuomas Leinonen

Työn nimi: Keilaavan millimetriaaltoradiolinkin suuntaaminen ja seuraaminen

Päivämäärä: 23. syyskuuta 2015 Kieli: Englanti Sivumäärä: 15+136

Radiotieteen ja -tekniikan laitos

Professuuri: Radiotiede ja -tekniikka

Työn valvoja: Prof. Ville Viikari

Työn ohjaaja: DI Jyri Putkonen

Seuraavan sukupolven matkaviestinjärjestelmien erittäin nopeissa datayhteyksissä tullaan hyödyntämään millimetriaaltoteknologiaa. Näillä taajuuksilla radioympäristö on kuitenkin hyvin haastava, mikä edellyttää verkon solutiheyden moninkertaistamista, täysin langattomia tukiasemia ja erittäin suuntaavia antenneja. Tässä diplomityössä tutkitaan eri keinoja kuinka tukiasemien väliset linkit kohdistetaan tehokkaasti, ja miten se vaikuttaa verkon rakenteeseen ja hallintaan.

Työ tarjoaa kattavan taustaselvityksen mm-aaltosoluverkon toteuttamiseen tarvittavista asioista. Keilanohjausta tarkastellaan sekä verkon automaattisen laajentamisen että kohteen aktiivisen seurauksen kannalta. Tämän lisäksi työssä tutkitaan keilattavia antenneja, mm-aaltojen etenemistä kaupunkiympäristöissä ja verkkorakennetta. Näiden lisäksi työssä rajataan edellytykset, esitetään mahdollisia ratkaisuja, ja vertaillaan näitä olemassa oleviin toteutuksiin.

Nykyiset keilaustoteutukset ovat pääasiassa suunniteltu lyhyen kantaman sisäyhteyksille, eivätkä siten vastaa ongelman asettelua. Aikaisempi toiminnallisuus keskittyy yhteen ainoaan linkkiin vaikka tehokas toteutus huomioisi koko järjestelmän kohdistusongelman fysikaalista perustaa unohtamatta. Verkkolaitteiden tulisi hyödyntää olemassa olevaa radioverkkoa, sekä paikka- että suuntatietoja, ja koneoppimisen keinoja. Vaikka aiheeseen liittyvä teknologia on kehittynyt viime vuosina harppauksin, mm-aaltosoluverkot ovat kaikkea muuta kuin valmiita markkinoille.

Avainsanat: 5G, algoritmi, antenni, keilanmuodostus, keilanohjaus, millimetriaalto, radio, radioaaltojen eteneminen

Preface

This thesis was written while I worked for Nokia Networks[1] in Espoo, Finland between April and September 2015. The work is related to a European collaborative project called *MiWaveS* that aims to develop and promote millimeter-wave wireless communication technologies in upcoming mobile networks. More information on the project may be found by following the link `http://www.miwaves.eu/`.

I would like to thank Nokia for giving me this wonderful opportunity to work on the forefront of next-generation mobile networks. I would also like to thank both my supervisor Prof. Ville Viikari and my advisor Jyri Putkonen, M. Sc. (Tech.) for their invaluable guidance and feedback throughout the process. There are also a number of other people who deserve my highest commendation for their support and suggestions, including Risto, Ola, Anu, Luukas and the rest of my peer reviewers.

As I am about to complete my studies, I find myself reflecting on the events and all the people I have met over the past six years. All those joyful memories bring tears to my my eyes. Thus, I am signing off by offering my most sincere thanks and deepest bows to my family and my friends all over the world for their support and understanding in the course of my studies. You made the difference by making this journey worthwhile. This thesis is for you!

Espoo, Finland
September 23rd, 2015

Tuomas Leinonen

[1] `https://networks.nokia.com/`

Contents

Symbols, operators and abbreviations

Symbols

Scalars

a	(positive) real number (used in function definitions)
A_{eff}	effective antenna area
b	(positive) real number (used in function definitions)
C	scan coverage
c	the speed of light in used transport medium
D	antenna displacement
d	inter-element spacing
d_i	spacing between the first and ith antenna element
e	error in (estimated) position
$\ddot{e}$	error in (estimated) acceleration
E	antenna excitation amplitude
f	frequency
F_N	radius of the N^{th} Fresnel ellipsoid ($N \in \mathbb{N}^+$)
G	gain
G_{c}	co-polarization gain
G_{T}	transmit antenna gain
G_{R}	receive antenna gain
G_{x}	cross-polarization gain
$k_{n,\,l}$	phase-shift exponent for nth antenna element and lth beam
k	wave number
K	(positive) natural number (count)
n	(positive) natural number (index)
L	attenuation
L_{atm}	attenuation due to atmospheric gases
L_{pol}	polarization mismatch loss
L_{rain}	attenuation due to rainfall
m	integer (phase compensation)
M	(positive) natural number (count)
M_{R}	the number of receive antennae

M_T	the number of transmit antennae
n	(positive) natural number (index)
n_PL	path loss exponent
N	(positive) natural number (count)
N_R	the number of parallel receive chains
N_T	the number of parallel transmit chains
P	power
P_C	power of the co-polarized component
P_T	transmit power (delivered to the antenna)
P_R	received power (available from the antenna)
P_X	power of the cross-polarized component
r	link length
r_1	distance to the POI from one endpoint
r_2	distance to the POI from the other endpoint
R	radius of the half-power illumination disc
s	position
$\dot{s}$	velocity
$\ddot{s}$	acceleration
t	time
α	rain attenuation power-law exponent
α_x	roll angle around the longitudinal x axis
β_y	yaw angle around the vertical y axis
γ_z	pitch angle around the lateral z axis
Δf	frequency difference or bandwidth
$\Delta\phi$	phase difference
Δt	time delay
Δx	distance on the x axis
θ	incident angle, angle of arrival
ϑ	elevation angle
ϑ_R	elevation angle at the receiver
ϑ_T	elevation angle at the transmitter
κ	rain attenuation power-law coefficient

λ	wavelength
ρ	precipitation rate
σ_{SF}	shadow fading loss
τ	polarization tilt angle
ϕ	phase shift
φ	azimuth angle
φ_{R}	azimuth angle at the receiver
φ_{T}	azimuth angle at the transmitter

Arrays

In this thesis, **upright boldface** characters designate arrays. Lowercase letters like **a** are vectors whereas capital letters like **A** are matrices.

$\mathbf{H}$	(dual-polarized) channel matrix
$\mathbf{H}_{\mathrm{HH}}$	horizontal to horizontal polarization channel matrix
$\mathbf{H}_{\mathrm{HV}}$	vertical to horizontal polarization channel matrix
$\mathbf{H}_{\mathrm{VH}}$	horizontal to vertical polarization channel matrix
$\mathbf{H}_{\mathrm{VV}}$	vertical to vertical polarization channel matrix
$\mathbf{I}$	identity matrix
$\mathbf{I}_2$	two-by-two identity matrix
$\mathbf{N}$	noise and interference matrix
$\mathbf{n}$	noise and interference vector
$\mathbf{n}_{\mathrm{H}}$	horizontally polarized noise and interference vector
$\mathbf{n}_{\mathrm{V}}$	vertically polarized noise and interference vector
$\mathbf{S}$	original signal matrix
$\mathbf{s}$	original signal vector
$\mathbf{s}_{\mathrm{H}}$	original horizontally polarized signal vector
$\mathbf{s}_{\mathrm{V}}$	original vertically polarized signal vector
$\mathbf{V}$	baseband weight matrix
$\mathbf{V}_{\mathrm{H}}$	baseband weight matrix for horizontally polarized signals
$\mathbf{V}_{\mathrm{R}}$	baseband combiner weight matrix
$\mathbf{V}_{\mathrm{T}}$	baseband precoder weight matrix
$\mathbf{V}_{\mathrm{V}}$	baseband weight matrix for vertically polarized signals
$\mathbf{V}_{\mathrm{R,V}}$	baseband combiner matrix for vertically polarized signals

$\mathbf{V}_{\mathrm{R,H}}$	baseband combiner matrix for horizontally polarized signals
$\mathbf{V}_{\mathrm{T,V}}$	baseband precoder matrix for vertically polarized signals
$\mathbf{V}_{\mathrm{T,H}}$	baseband precoder matrix for horizontally polarized signals
$\mathbf{W}$	antenna weight matrix
$\mathbf{W}_{\mathrm{H}}$	antenna weight matrix for horizontally polarized signals
$\mathbf{W}_{\mathrm{R}}$	receive antenna weight matrix
$\mathbf{W}_{\mathrm{T}}$	transmit antenna weight matrix
$\mathbf{W}_{\mathrm{V}}$	antenna weight matrix for vertically polarized signals
$\mathbf{W}_{\mathrm{R,H}}$	receive antenna matrix for horizontally polarized signals
$\mathbf{W}_{\mathrm{R,V}}$	receive antenna matrix for vertically polarized signals
$\mathbf{W}_{\mathrm{T,H}}$	transmit antenna matrix for horizontally polarized signals
$\mathbf{W}_{\mathrm{T,V}}$	transmit antenna matrix for vertically polarized signals
$\mathbf{X}$	received signal matrix
$\mathbf{x}$	received signal vector
$\mathbf{x}_{\mathrm{H}}$	received horizontally polarized signal vector
$\mathbf{x}_{\mathrm{V}}$	received vertically polarized signal vector

Other

$\vec{a}$	vector (used in function definitions)
$\vec{b}$	vector (used in function definitions)
$\vec{E}$	electric field vector
$\vec{H}$	magnetic field vector
$\vec{p}$	polarization vector
$\vec{p}_{\mathrm{R}}$	receive polarization vector
$\vec{p}_{\mathrm{T}}$	transmit polarization vector
$\vec{S}$	Poynting vector
$\vec{x}$	unit vector toward the positive x axis
$\vec{y}$	unit vector toward the positive y axis
$\vec{z}$	unit vector toward the positive z axis
$\mathcal{A}$	(unordered) set (used in function definitions)
$\mathcal{B}$	(unordered) set (used in function definitions)
$\mathbb{Z}$	the set of all integers

$\mathbb{N}$	the set of all natural numbers
$\mathbb{N}^+$	the set of all positive natural numbers
$\mathbb{R}$	the set of all real numbers
$\mathbb{R}^+$	the set of all positive real numbers
j	imaginary unit ($\mathrm{j} = \sqrt{-1} = 1\angle 90\,^\circ$)
i	loop variable or index
j	loop variable or index
$\mathcal{I}$	probing sequence
$\mathcal{J}$	probing sequence
$\mathcal{S}$	ordered set, i.e., sequence
ν	substitute integral variable for time t
υ	substitute integral variable for time t

Operators and functions

$\lvert\vec{a}\rvert$	magnitude of the vector $\vec{a}$
$\vec{a} \times \vec{b}$	cross product of vectors $\vec{a}$ and $\vec{b}$
$\vec{a} \cdot \vec{b}$	dot product of vectors $\vec{a}$ and $\vec{b}$
$\mathbf{A}^{\mathrm{T}}$	transpose of matrix $\mathbf{A}$
$\mathbf{A}^{\mathrm{H}}$	conjugate transpose of matrix $\mathbf{A}$
$\mathbf{A} \odot \mathbf{B}$	Hadamard product of matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A}\,\mathbf{B}$	product of matrices $\mathbf{A}$ and $\mathbf{B}$
$a \in \mathcal{A}$	element a is a member of set $\mathcal{A}$
$\mathcal{A} \setminus \mathcal{B}$	set $\mathcal{A}$ excluding elements also in set $\mathcal{B}$
$f(n) \in \mathcal{O}\{g(n)\}$	$\exists\, a \in \mathbb{R}^+$ and $N \in \mathbb{N}$ such that $f(n) < ag(n), \forall n > N$
$\lfloor a \rfloor$	round toward zero, i.e., floor
$\mathrm{mod}(N, M)$	remainder of N divided by M where $N, M \in \mathbb{N}^+$
$\mathrm{lcm}(N, M)$	least common multiple of N and M where $N, M \in \mathbb{N}^+$

$\text{rep}(\mathcal{S}, N)$	sequence $\mathcal{S}$ is repeated N times where $N \in \mathbb{N}^+$
$\text{shuffle}(\mathcal{S}, N)$	sequence $\mathcal{S}$ is independently shuffled and repeated N times where $N \in \mathbb{N}^+$

Abbreviations

3D	three-dimensional
3GPP	3rd Generation Partnership Project
4G	fourth generation mobile networks
5G	fifth generation heterogeneous mobile networks
ABF	analog beamforming or beamformer
ADC	analog-to-digital converter
AGC	automatic gain control
AL	access link
ANT	antenna
AoD	angle of arrival
AoD	angle of departure
AP	access point
BB	baseband
BF	beamforming
BH	backhaul
BoM	bill of materials
BPSK	binary phase-shift keying
BS	base station
CCDP	co-channel dual polarization
CCI	co-channel interference
CDMA	code-division multiple access
CSI	channel state information
CSMA	carrier sense multiple access
CTRL	control circuit or signal
DAC	digital-to-analog converter
DBF	digital beamforming or beamformer
DFT	discrete Fourier transform
D-MPS	mode pursuing sampling method for a discrete variable space

DoA	direction of arrival
DSP	digital signal processing
EIRP	equivalent isotropic radiated power
EM	electromagnetic
ESPRIT	Estimation of Signal Parameters via Rotational Invariant Techniques
ETSI	European Telecommunications Standards Institute
FDD	frequency-division duplex
FDE	frequency domain equalization
FDMA	frequency-division multiple access
FFT	fast Fourier transform
FPGA	field-programmable gate array
FSPL	free-space path loss
GEN	signal generator
GNSS	global navigation satellite system
GSCM	geometry-based stochastic channel model
GW	gateway
HBF	hybrid beamforming or beamformer
HC	host controller
HetNet	heterogeneous network
HPBW	half-power beamwidth
I2I	indoor-to-indoor
ID	identifier
IEEE	Institute of Electrical and Electronics Engineers and Electronics Engineers
IF	intermediate frequency
IFFT	inverse FFT
ILSP-CMA	Iterative Least Squares Projection Based Constant Modulus Algorithm
IMT	International Mobile Telecommunications
I/O	input / output
IP	interaction point
I/Q	in-phase / quadrature-phase
ISI	intersymbol interference

ITU	International Telecommunication Union
ITU-R	ITU Radiocommunication Sector
LOS	line of sight
LTE	Long-Term Evolution
MBB	mobile broadband
METIS	Mobile and wireless communications Enablers for the Twenty-twenty Information Society
MIMO	multiple input, multiple output
mmW	millimeter wave
MSC	multiple sidelobe canceler
MU-MIMO	multi-user MIMO
MUSIC	MUltiple SIgnal Classification
NGMN	next-generation mobile networks
NI	National Instruments
NLOS	non-line of sight
NMT	Nordisk Mobiltelefoni
O2I	outdoor-to-indoor
O2O	outdoor-to-outdoor
OFDM	orthogonal frequency division multiplexing
OS	orientation sensor
P2MP	point-to-multipoint
P2P	point-to-point
PAPR	peak-to-average power ratio
PC	personal computer
PCI	Peripheral Component Interconnect
PL	(omnidirectional) path loss
POI	point of interest
PS	phase shifter
PSK	phase-shift keying
PXI	PCI eXtensions for Instrumentation
QAM	quadrature amplitude modulation
QoE	quality of experience
QoS	quality of service

REF	reference
RF	radio frequency
RNG	random number generation or -tor
RSSI	received signal strength indicator
RX	receiver
SC	single-carrier
SCM	spatial channel model
SDMA	space-division multiple access
SVD	single value decomposition
SFFT	spatial FFT
SINR	signal-to-interference and noise-ratio
SU-MIMO	single-user MIMO
SNR	signal-to-noise-ratio
SOHN	self-organizing heterogeneous network
SON	self-organizing network
TCO	total cost of ownership
TDD	time-division duplex
TDMA	time-division multiple access
TRX	transceiver
TTL	time to live
TX	transmitter
UE	user equipment
UCA	uniform circular array
ULA	uniform linear array
UMa	urban macro
UMi	urban micro
UPA	uniform planar array
UTD	uniform theory of diffraction
VGA	variable gain amplifier
WINNER	Wireless World Initiative New Radio
WLAN	wireless local area network
WPAN	wireless personal area network
XPD	cross-polarization discrimination
XPIC	cross-polarized interference canceler

1 Introduction

Mobile communications and smart devices have been the fastest growing consumer electronics segment in the past decade, and there is no end in sight [1]–[3]. Instead, the dominance has only strengthened as consumers are taking full advantage of smart applications enabled through seamless connectivity [4]. Consumers simply want and expect more; to be completely immersed in the possibilities of the connected world.

Mobile broadband (MBB) data volumes are expected to continue increasing exponentially far into the future [5]–[10]. Cisco [5] and Ericsson [6] report the compound annual growth rate (CAGR) for 2014 to have been more than +60 %. Future projections range from eight-fold increase over six years [6] to ten-fold increase over five years [5]. These forecasts correspond to a CAGR of +40 % and +60 %, respectively.

One of the main driving forces behind the growth is ultra high-resolution video [5], [6], [8]–[12], expected to generate more than half of the annual mobile traffic [5], [6]. Some other significant sources of traffic are cloud computing, entertainment, the simultaneous use of multiple smart devices such as tablets and wearables, and applications we cannot even fathom at this moment [6], [8]–[11]. Machine-to-machine (M2M) communications and the Internet of Things (IoT) are also expected to cause considerable increase in mobile traffic in the future, amounting to more and more over time [5], [6], [8]–[10], [13], [14].

Faced with an exponential increase in both the number of users and data traffic, current mobile networks cannot simply live up to such high demands and consumers' expectations [14]–[16]. Consequently, all major networking equipment providers [6], [14], [17], [18], operators [10], standardization bodies and other consortiums alike [8], [9], [16], [19], [20] focus on developing next-generation mobile networks (NGMN) and associated services. Moreover, the concept of a connected world has attracted some new parties, even governments [21]–[26]. Inter-vehicle communications and intelligent transportation systems [8]–[10], [26] have for instance caught the attention of the entire automotive industry.

Reports by parties involved in NGMN research [6], [10], [14], [16]–[18], [20]–[24] provide some insight into what kind of goals and requirements are set for the fifth-generation mobile networks (5G). Targeting for widespread commercial deployment in early 2020's, some of the envisioned goals are a 1000-fold increase in capacity, peak data rates in excess of 10 Gbps and end-to-end latencies under 1 ms. Furthermore, improvements in energy efficiency, reliability, cost-effectiveness and limiting exposure are also of top concern.

While actual 5G standardization is still in its infancy and only taking its first steps, there is a clear consensus among different parties concerning the major outlines [6], [10], [14], [16]–[18], [20]–[24]. Unlike existing cellular networks, 5G networks will be *scenario driven* to reflect the shift from human-centric communications toward co-operative ubiquitous computing. In the 2020's information society, Internet access

is available anywhere, anytime, to anyone and to anything – seamlessly even on the go.

Different applications and use cases have very different goals and requirements, and trying to cover everything within this broad scope at once would not be possible. All the aforementioned goals need not be met simultaneously. Instead, the network will adapt to current situation and optimize the key performance metrics to provide ultimate user experience or service quality. For instance, high throughput is of the essence in large data transfers, while massive sensor networks should focus on energy efficiency. Further, real-time systems emphasize responsiveness and availability, requiring extremely low latency and utmost reliability. Other visioned scenarios include high device mobility, temporary mass gatherings and efficient disaster relief, to name a few examples.

The overall network performance must thus improve drastically on a number of fronts, should these requirements be met. In light of today's typical use cases, the key target is to increase both capacity and throughput. Three means regularly are underlined as key enablers: more spectrum (Hz), more efficient use of spectral resources (bits/s/Hz/m^2) and massive densification of *small cells* [7], [27], [28]. Massive densification enables a new frontier of spatial processing which is typically seen as a gateway to substantial gains from the first two means [16], [28]–[30].

Small cells are low-power access points (AP) with limited range of roughly up to a hundred meters. They are used either to extend high-speed coverage or more commonly to increase network capacity [30]–[32]. Capacity is improved by offloading traffic from macro base stations (BS) to individual APs in crowded areas. Such high-demand locations include targets such as food courts in shopping malls and public transportation hubs such as metro stations.

Small cells are already being deployed as a part of existing networks [30]–[32], and a similar architecture is envisioned to play a key role in the future [7], [12], [28], [33]. This heterogeneous architecture is shown in Figure 1 where a BS installed on a rooftop provides macro coverage supplemented by a number of street-level small cell APs connected in a wireless mesh [29], [32], [34]–[36]. Given the envisioned requirements for 5G and the crowded lower end of the radio spectrum, these small cells will use *millimeter-wave* (mmW) frequencies [14], [28], [33], [34], [37]. It is also possible for these APs to provide connectivity using other wireless technologies such as Wireless Local Area Networks (WLAN). Numerous operators are already offering so called Wi-Fi hotspots scattered in urban areas based on the concept [32], [34].

The deployment of mmWs is motivated by the numerous favorable characteristics they exhibit [28], [32], [34], [38], [39]. First, there are many gigahertz of unlicensed and light-licensed spectrum available globally in the V- (57 ... 66 GHz) and E-bands (71 ... 76 GHz and 81 ... 86 GHz), respectively, enabling high throughput. Second, short wavelength makes small radio frequency (RF) equipment. Third, high attenuation reduces interference allowing superior frequency reuse through spatial separation. Fourth, multicarrier division is typically unnecessary leading to reduced

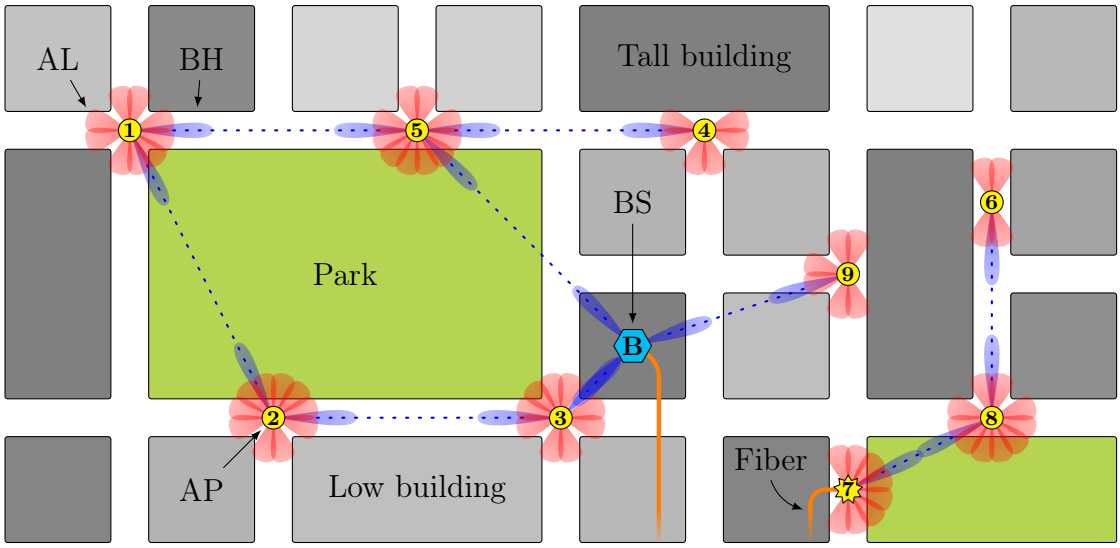


Figure 1: Envisioned fifth-generation heterogeneous mobile network architecture.

hardware complexity and energy consumption. In fact, mmWs are already employed for short-range indoor communications using the 60 GHz band [40]–[43].

There are, however, a number of challenges to be addressed before millimeter waves may be effectively deployed in mobile communications. These are related to the unfavorable radio environment at higher frequencies [32], [38], [44]. High free-space path loss, atmospheric losses and weather effects have a dramatic effect on link range and quality. Furthermore, the radio channel behavior is altogether different from that at conventional frequencies since *all* dimensions are effectively 10 ... 100 times larger [45]–[48]. These drawbacks are addressed by limiting the required cell range by increasing their number, and by using high-gain antennas at both ends of the link [32], [38], [44]. The links are highly directional in nature, and any obstacles in the line-of-sight (LOS) path can easily render the link unusable [32], [38], [44], [49].

These in turn create a new set of challenges. First, massive densification warrants that the mmW APs be installed close to the user, as is depicted in Figure 1 [32], [34], [35], [44]. The APs will be installed on building facades and on street furniture such as lamp posts, and in other unconventional locations where wired connectivity is not practical. Millimeter-wave links are thus needed to provide connectivity not only to APs but also further to the users. Furthermore, these mounting structures may be physically unstable, prone to vibration, and may sway under heavy winds [13], [21], [30], [32], [44], [50]–[52].

Second, high-gain antennas are also very directive, meaning they have narrow pencil-like beams only couple of degrees in width, requiring precise alignment [32], [34], [38], [53], [54]. This concerns both access links (AL) connecting the moving user to the swaying AP and backhaul links (BH) connecting vibrating APs to each other and the macro BS. Both scenarios require antennas with electronically steerable beams.

Antenna alignment is crucial for making the envisioned architecture feasible not only connectivity-wise but also financially. Substantial investments are required from the operators with every new network generation – now more than ever thanks to massive densification [7], [29], [32], [34], [36]. Consequently, minimizing all costs associated with designing, installing, operating and maintaining these numerous small cells is of vital importance. Spatial separation and minimal interference also make it possible to avoid extensive network design, planning and optimization beforehand [29], [34]–[36].

The total cost of ownership (TCO) is minimized through automation, intelligence, agility, flexibility and support for incremental approach [10], [11], [16], [29], [35], [36]. The networking equipment should work in a *plug and play* fashion without the need for manual configuration such as antenna alignment. Further, the network should take full advantage of cloud processing while exhibiting cognitive features, such as organizing itself to optimize both quality of service or experience (QoS or QoE, respectively), and energy efficiency. Antenna alignment plays a central role in achieving all these goals.

This Master's thesis investigates the possibility of using millimeter waves to provide high-throughput Internet access in a dense urban information society. The work focuses on the beam steering functionality necessary for wireless backhaul using millimeter waves. Targeting a system similar to what was introduced above, the work considers an in-depth analysis of the necessary functionality, namely automatic link alignment and tracking, and their interplay with the surrounding system. The work further specifies requirements, proposes possible approaches and compares those with existing solutions. Electronically steerable antennas and mmW propagation are also thoroughly discussed.

This document is organized in the following manner. In Section 2, different beam steering antennas and techniques are discussed. Section 3 concerns the radio wave propagation at mmW frequencies and channel modeling. Section 4 considers the envisioned network architecture and how the architecture affects beam alignment. Sections 5 and 6 form the core of the thesis and present author's original contribution. The former section covers network expansion whereas the latter considers maintaining existing links. An overview of existing beam steering algorithms are given in Section 7. Finally, Section 8 concludes the outcome of this work.

2 Electronic beam steering

This section considers the foundations of beam steering while radio wave propagation at millimeter-wave frequencies is studied in Section 3. The discussion on beam steering covers the basic principles, different mechanisms and their comparison. Emphasizing hardware aspects, examples of practical implementations of steerable antennas are introduced.

Before discussing beam alignment, the associated terminology is introduced. The goal is to differentiate between various similar and possibly overlapping definitions to help the reader avoid confusion. The following terminology is considered to be the most coherent, and will be used throughout the thesis. It is also preferred in the industry, even though different terminology can occasionally be seen.

First, we make a distinction between alignment and other concepts. *Antenna tuning* refers to changing the operating frequency of an antenna for instance by tuning its matching [55]. A somewhat similar notion, *antenna reconfiguration*, refers to a wide range of possible adaptive or adjustable parameters such as impedance, radiation pattern, polarization, or a combination of these [55]. Antenna tuning is a common way to make antennas reconfigurable.

Considering alignment, *antenna alignment* refers to directing the radiation from an antenna toward its target [55]. The antenna may be directed either manually or automatically, statically or dynamically, or electronically or mechanically. Static and dynamic alignment are commonly referred to as *antenna pointing* and *antenna tracking*, respectively [55]. This work considers electronic antenna alignment done automatically without user or operator intervention. Our use case includes both static and dynamic tasks, namely new link establishment and maintaining existing links, respectively.

Focusing the antenna main beam electronically on the intended target is referred to as *beam steering*. In addition to rotating or moving the antenna (feed) mechanically, there are two common beam steering mechanisms as follows [54]–[58]. It is possible to switch between different feed configurations or perform *beamforming* (BF) using an antenna array. A third possibility is to use a tunable focusing element such as a reflectarray or an electronic lens. However, both are based on the beamforming concept.

The focus of the following discussion is on beamforming due to its superior versatility. This is also the reason why beamforming is often associated with *smart* or *adaptive antenna systems*. Using sophisticated digital processing, a radio system may adapt to its environment by simultaneously directing radiation maxima and minima toward desired sources and interferers, respectively [54]–[59]. In addition to excellent performance, antenna arrays are preferred over other structures given their small form factor.

Beamforming is not a new concept since applications exist especially in the context of radars and satellite communications [55], [56], [58], [60], [61]. Beam steering is also included in a number of recent standards intended for short-distance indoor

communications using the unlicensed 60 GHz frequency band. Maintained by the Institute of Electrical and Electronics Engineers (IEEE), the most famous of these standards are the IEEE 802.11ad [40] and IEEE 802.15.3c [41] for wireless local and personal area networking (WLAN and WPAN), respectively. Other 60 GHz standards include ECMA-387 [42] and WirelessHD [43], intended for wireless high-quality video streaming.

These standards are further discussed in Section 7.2 where a number of existing beam steering implementations are showcased. This section covers the basics, starting with beamforming theory in the following subsection. We discuss the physical phenomena, possibilities and limitations. This is followed by discussion on different beamformer implementations in Section 2.2. Other steering techniques are studied in the last subsection.

2.1 Fundamentals of beamforming

The directivity of an antenna is related to its electrical size with larger antennas having more gain. Instead of having one large element, an antenna may also be made more directive using multiple smaller elements placed close to each other [54]–[56], [58], [62]. Feeding these elements simultaneously constitutes an *antenna array*. Arrays are preferred for their steerability and other benefits, achieved at an acceptable cost of increased feed network complexity. In practice, most arrays are *uniform*, i.e., comprise similar equispaced elements for easier analysis and synthesis. The following discussion is limited to such arrays.

Antenna arrays operate based on interference, and steerability is achieved by controlling the directions in which waves interfere constructively and destructively [54]–[56], [58], [62]. Take the situation pictured in Figure 2, where an antenna array comprising two similar sources is investigated. The elements are placed distance d apart, and are equal in amplitude E but differ ϕ in phase. Interested in the far-field behavior at a distance $r, r \gg d$ the radiated electric field $\vec{E}$ may be assumed to be a plane wave proportional to

$$\vec{E} \propto \frac{e^{-\mathrm{j}kr}}{r}, \tag{1}$$

where $k = 2\pi/\lambda$ is the wave number and $\lambda = c/f$ the wavelength [54]–[56], [58], [62]. The wavelength depends on the transmit medium, found as the ratio of speed of light in said medium c (in air $c \approx 2.998 \cdot 10^8$ m/s) and frequency f.

Since $r \gg d$ at the point of interest (POI), both the incident angle θ and the electric field magnitude $|\vec{E}|$ may be assumed equal for both sources [54]–[56], [58], [62]. The difference in distance between the elements d is thus meaningful only in phase. The waves interfere constructively or destructively when they are *in phase* and *out of phase*, respectively. For a plane wave at incident angle θ, the difference in distance between the two elements is $d \sin\theta$. For *constructive* interference, the phase difference caused by this distance should be compensated by the phase shift ϕ. That

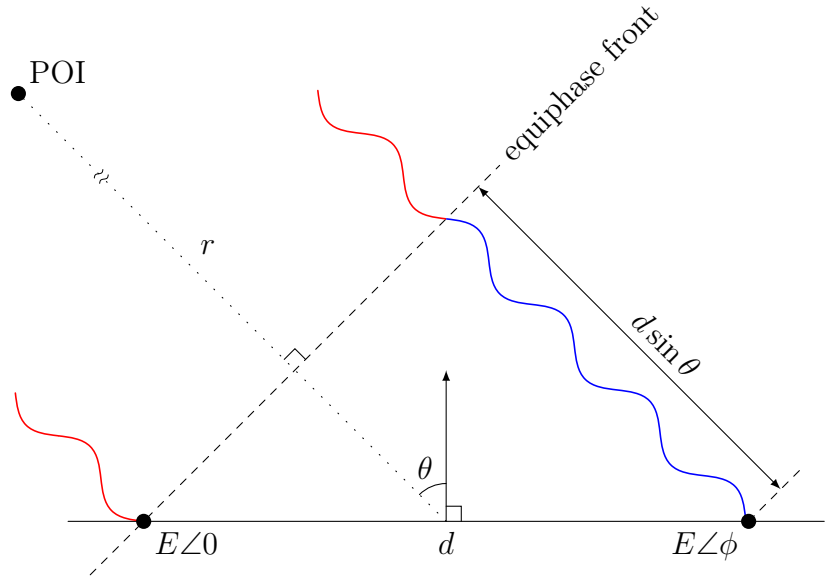


Figure 2: Far-field ($r \gg d$) interference of a two-element array of similar sources.

is,

$$\phi - kd \sin\theta = m2\pi, \tag{2}$$

where m is an integer, i.e., $m \in \mathbb{Z}$. *Destructive* interference occurs when the left side equals to $(2m + 1)\pi$. In Figure 2, the signals are out of phase and interfere destructively ($m = 3$).

The same phenomenon may also be used to determine the *angle of arrival* (AoA), also referred to as *direction of arrival* (DoA), used in smart antenna systems to steer the beam toward the source [55]–[59]. Using multiple receiver chains, the relative time delay Δt between the signals in different chains can be measured. The angle of arrival θ is then found as follows

$$\theta = \arcsin \frac{\Delta t c}{d}. \tag{3}$$

The transmitter is analogous; the angle of departure (AoD) may be set reciprocally by controlling the delays between different transmitter chains. This is the foundation of digital beamforming discussed further in Section 2.2.

Previous results may be generalized to arrays with arbitrary number of elements in more than one dimension, and to arrays of different geometry [54]–[56], [58], [62]. For instance, uniform circular arrays (UCA) benefit from symmetry, and the beam maintains its shape at all steering angels. Further, the total radiation pattern of an array comprising similar elements may be found as the product of the radiation pattern of a single element and the array factor [54]–[56], [58], [62]. The array factor is determined by the geometry and the *complex weighting*, i.e., the excitation, of individual elements. The antenna radiation pattern is thus controlled by altering the array factor, or the weighting to be more specific.

In the analysis above, only phase shift was used to implement simple beam steering. A single rotated beam is created when the phases of individual elements are set progressively so that the condition in Equation (2) is met for all elements i (i is a positive natural number: $i \in \mathbb{N}^+$) with different distances d_i respective to the first element [54]–[56], [58], [60], [62]. The width of this beam is determined by the number of elements, their characteristics and the steering angle. Exact one-to-one rotation is not possible and the pattern is distorted with large steering angles. While the fundamental cause is the array geometry, the effect is more significant the more directive the individual elements are.

Progressive phase shifting does not provide variable beamwidth or sidelobe minimization by itself. That would require control over amplitude and windowing techniques. Simple windowing may be achieved using 1-bit On/Off switching, while more precise control requires variable gain amplifiers (VGAs) or tunable attenuators. However, rudimentary beamwidth control is still possible by superimposing these simple beams. This is achieved by dividing the array into multiple subarrays, and directing each group separately. The subarrays are treated as individual elements, and the final array is created by grouping these subarrays together. Connected to the same transceiver, the beams carry the same data and all the previous principles apply [54]–[56], [58], [60], [62].

Without windowing, this approach offers only limited control over the beamwidth before the beams become distinct. Even with small angular offsets, the resulting pattern is contains multiple local minima and maxima in the main beam. Moreover, small offsets require accurate phase shifters that increase complexity, bill of materials (BoM) and power consumption. This idea is elaborated in Figure 3 showing radiated electric field magnitude of a phased antenna array in two different configurations. Both subarray beams are shown in blue and green in addition to the red sum beam. The array is divided into two subarrays, both of which are assumed to follow a sinc^2 radiation pattern with a half-power beamwidth (HPBW) and electric field amplitude of $10\,^\circ$ and 1, respectively.

On the left, both subarrays have equal phasing resulting in a $10\,^\circ$ beam with amplitude of 2. On the right, the subarrays are phased to create two slightly offset beams – a process referred to as *squinting* or *beam spoiling* [52], [63], [64]. The resulting beam has more than twice the beamwidth but roughly only half the electric field amplitude – note the axes. The HPBW may thus be more than doubled by sacrificing 6 dB from gain while also tolerating 1 dB gain ripple.

More versatile pattern generation is possible if both the phase and amplitude of each individual element are controlled [54]–[56], [58]. In addition to being able to control the beamwidth and sidelobes levels, we may use multiple beams and place nulls toward multiple interferers. Still, uniform weighting is favored for its lower complexity and limited power consumption, and such systems are commonly referred to as *phased antenna arrays* [54]–[56], [58], [60], [62]. Moreover, the phase control is

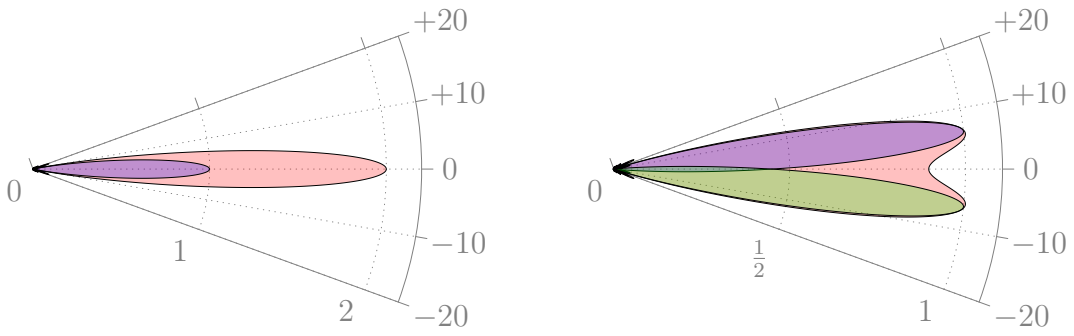


Figure 3: Beamwidth control using two phase-offset sinc^2 beams.

often implemented using a number of predetermined, enumerated values taken from a codebook [52], [65]–[71].

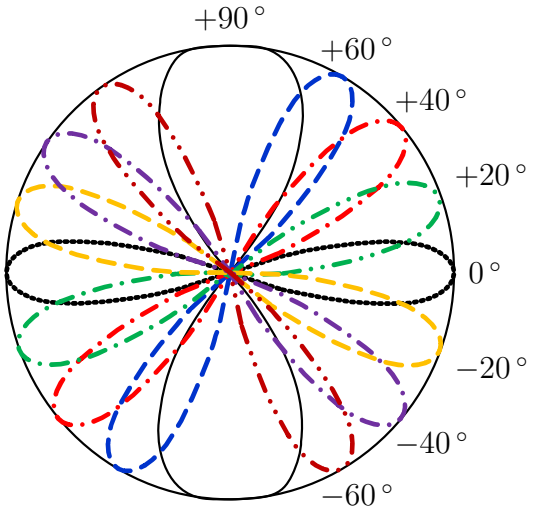
A *codebook* is table that maps a number of predetermined beams to sets of antenna excitation values [52], [65]–[71]. A single set determines the weights for each element corresponding to a certain beam. Such a collection of weights is commonly referred to as *a beamforming vector*, *an antenna weight vector* or *a beam steering vector*. The terms are interchangeable and refer to the same exact thing.

The codebook design concerns finding a suitable array factor. Starting with the array factor expression, a custom codebook may be generated for instance by using *genetic algorithms* [72]–[74]. For rigorous analysis, an electromagnetic simulator could be used. While this is considerably more tedious, simulations can also account for adverse effects such as mutual coupling and array (mis)calibration [55]–[57].

However, there are a number of well-known codebook families [65], [67], [70], [75] open for use. For instance, the ECMA-387 standard [42] defines both Fourier and Hadamard codebooks for half-wave spaced linear arrays based on discrete Fourier transform (DFT) and Hadamard matrices, respectively. In general, the phase shifting values are obtained using a mathematical formula as a function of inter-element spacing, phase shifting range and accuracy, and desired number of beams.

If there are no constraints regarding the phase shift, the codebook may be found using progressive phasing. In practice, however, it is important to consider both the phase shifting accuracy and its limits. The elements are typically spaced $\lambda/2$ apart and the phase shifters support values $\{0, 1, \ldots, 2^N - 1\} \cdot 360\,^\circ / 2^N$ using N bits ($N \in \mathbb{N}^+$). In practice, codebooks are usually rather small to limit both the search area (see Section 5) and the required phase shifting accuracy.

Similarly to the ECMA-387 standard, the IEEE 802.15.3c 60 GHz WPAN standard defines a standard codebook to be used with all implementations [41], [69], [75]. The simplest implementation comprises a four-element uniform linear array (ULA) and a codebook with eight approximately equispaced beams controlled by 2-bit phase shifters. The required *relative phase shift* for nth antenna element and lth beam



Beam	PS 1	PS 2	PS 3	PS 4
+90 °	0 °	180 °	0 °	180 °
+60 °	0 °	270 °	0 °	0 °
+40 °	0 °	270 °	90 °	90 °
+20 °	0 °	270 °	180 °	180 °
±0 °	0 °	0 °	0 °	0 °
−20 °	0 °	90 °	90 °	180 °
−40 °	0 °	90 °	180 °	270 °
−60 °	0 °	90 °	270 °	0 °

Figure 4: IEEE 802.15.3c codebook [41], [69], [75].

$\mathbf{W}(n, l)$ is found as

$$\mathbf{W}(n, l) = \mathrm{j}^{k_{n,\,l}}, \quad \text{where} \quad k_{n,\,l} = \left\lfloor \frac{4n}{M} \bmod \left(l + \frac{M}{2}, M \right) \right\rfloor . \tag{4}$$

Here, N and M $(N, M \in \mathbb{N}^+)$ are the total number of antennas and beams, respectively. The antennas and beams are one-indexed, i.e., $n \in \{1, 2, \ldots, N\}$ and $l \in \{1, 2, \ldots, M\}$. $\mathrm{j} = \sqrt{-1} = 1\angle 90\,^\circ$ is the imaginary unit, $\mathrm{mod}(N, M)$ is the remainder of N divided by M, and $\lfloor a \rfloor$ is the round toward zero, i.e., flooring function. The resulting codebook and beam patterns are shown in Figure 4. It is also possible to extend Equation (4) to support two-dimensional steering using uniform planar arrays (UPAs) [67].

It should be noted that in the discussion above, the real-life implementation is not considered. Broadly speaking, different beamforming approaches may be divided into three categories [28], [37], [67], [76]–[80]. First, there are pure-analog circuit-level solutions, referred to as *analog beamforming*. Second, one may also use highly advanced digital signal processing algorithms known as *digital beamforming*. Third, *hybrid beamforming* tries to combine these approaches to compromise between their pros and cons. These approaches are discussed in the next subsection.

2.2 Beamforming hardware

In analog beamforming (ABF), only one transceiver chain is used to feed all antennas [52], [54], [55], [58], [63], [67], [81]. At the receiving end, signals from different antennas are combined *after* complex weighting them in the analog domain using VGAs and phase shifters. This process is referred to as *precoding* or *antenna weighting*. The transmitter is analogous in structure; the excitation of individual elements is controlled using similar weighting scheme.

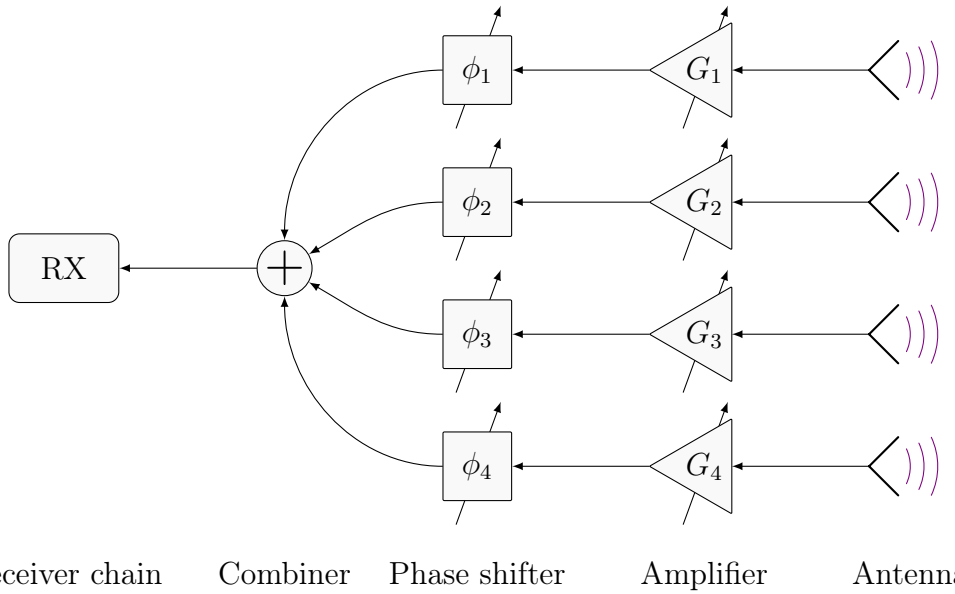


Figure 5: Analog beamforming receiver with four antennas and one receiver chain.

The structure described above is depicted in Figure 5 where a pure-analog beamforming receiver employing four antennas and one receiver chain (RX) is shown. In this example, true complex weighting is implemented using VGAs and phase shifters. Even though automatic gain control (AGC) is already used in virtually all wireless systems [82]–[85], precise element-wise gain control is typically not used but is still shown for completeness. The block diagram has been simplified otherwise, and a number of necessary stages are not shown. A complete receiver would include blocks for filtering, mixing, amplification and data conversion [53], [83]–[85].

In Figure 5, the complex weighting is carried out at radio frequency after mixing. Even though this is the most common approach in phased arrays, it is not the only option [81]. The individual signals may be weighted at intermediate frequency or at baseband before mixing, or even in the local oscillator (LO) path. As all these methods have their pros and cons, and none emerges as a clear winner, though Kong [81] argues for shifting at lower frequencies.

There are also some other drawbacks worth pointing out. These drawbacks are related to the phase shifting nature of the approach [54]–[56], [58], and from the non-idealities inherent in the extra components [53], [84]. One of these concerns is available bandwidth, since the steering functionality is based on relative phase difference. Referring back to Figure 2 and Equation (2), the phase difference $\Delta\phi$ caused by a bandwidth of Δf may be found as

$$\Delta\phi = \frac{2\pi}{c} d \sin\theta \, \Delta f. \tag{5}$$

This issue can be circumvented using frequency dependent phase shifters such as *switched delay lines* [53], [84]. As circuit area is often limited, delay lines are not

viable option at lower frequencies due to their size. Luckily, the circuit footprint of such designs is directly proportional to wavelength and thus decreases with increasing frequency.

Other concerns are phase accuracy and the shifting range [53], [58], [81], [84]. For non-periodic real-life signals carrying information over a finite bandwidth, the phase compensation condition is more strict than that shown in Equation (2). Instead of matching phases at one *point-frequency*, the signals need to be matched in *time* [58]. In other words, *all* frequencies within the used band at each moment in time should be matched to avoid aliasing. Consequently, the only value allowed for m is 0.

To further complicate the issue, the required phase shift (or corresponding delay) in progressive phasing may be greater than 2π rad or $360\,^\circ$. This is a problem as phase shifters usually exhibit modulo-2π behavior, meaning the maximum shift is 2π rad $= 360\,^\circ$. Luckily, the signal bandwidth is much lower than the carrier frequency, and the effect of both drawbacks is usually not significant for signal recovery. However, the antenna beams may be disturbed, resulting in lower gain or directional offset.

In digital beamforming (DBF), both the weighting and combining operations are carried out completely in the digital domain by means of digital signal processing (DSP) [55], [57]–[60], [67]. No element-wise phase shifters, VGAs or other analog components are used for precoding purposes. Instead, each antenna element has its own dedicated transceiver chain, and individual signals are digitized separately. A simplified block diagram of a four-way digital beamforming receiver is shown in Figure 6.

The greatest benefit of digital beamforming is the high degree of freedom and versatility due to multiple digitized data streams [55], [57]–[60], [67]. Essentially a

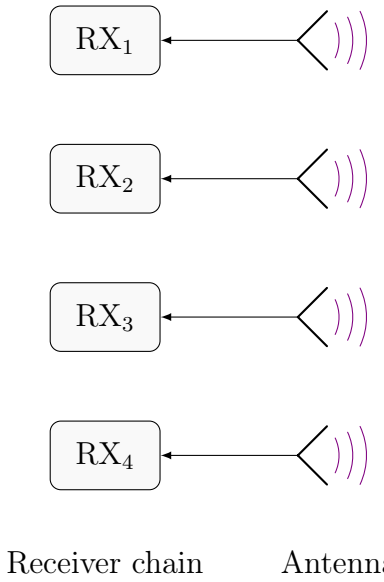


Figure 6: Digital beamforming receiver with four antennas and receiver chains.

system of simultaneous equations, these digitizer streams can be combined streams linearly to create or recover up to as many multiple independent data streams as there are digitizers. DBFs thus have native, simultaneous *point-to-multipoint* (P2MP) capability. In a multipath environment rich in fading, we may also use multiple streams to communicate with the same endpoint increasing throughput.

This space-time processing is referred to as *spatial multiplexing*, or more commonly as *multiple-input-multiple-output* (MIMO) [84], [86]–[90]. MIMO is possible if the signal to noise ratio (SNR) is high enough and the channel has multiple independent propagation paths due to interaction with objects in the environment – a topic discussed in greater depth in Section 3.4. If channel information is available at the receiver, these different data streams may be recovered. In some context, however, MIMO is used to refer to a system with a single transceivers but multiple antenna elements at both ends [63], [91]. This can be misleading since only a single data stream is used.

Combined with P2MP functionality, spatial multiplexing can simultaneously be used with multiple endpoints. These two different flavors of MIMO are referred to as single-user MIMO and multi-user MIMO (SU-MIMO and MU-MIMO), respectively. Furthermore, having multiple parallel transceivers helps to analyze the radio environment and to suppress interferers, resulting in an adaptive antenna system [54]–[59]

The greatest downfall of digital beamforming is its complexity [55], [57]–[60], [67], [80]. Not only some extra components are required, but the whole transceiver must be replicated for each antenna element. While MIMO processing is considered fundamental for modern wireless systems, only a few transceivers are required [84], [86]–[90]. Real beamforming systems, however, are typically much more complex, and DBF would increase both equipment cost and power consumption dramatically [55], [57]–[60], [67], [80]. In addition to aforementioned limitations, DBF requires high-performance DSPs and sophisticated algorithms. For these reasons, true digital beamforming has practically been limited to performance-critical applications such as radars and military communications [55], [58].

Compared to pure-analog approaches, digital beamforming offers superior performance and versatility but only at a cost too great for consumer applications. It is also possible to combine the approaches by grouping the elements into subarrays [28], [37], [67], [76]–[78], [80], [92]. Each subarray has its own transceiver and individual elements still have their own phase shifters and possibly VGAs for precoding. The system is simultaneously both a digital and an analog beamformer at subarray and element levels, respectively. Such a system is commonly referred to as a *hybrid beamformer* (HBF), and occasionally also as an *analog-digital*, a *mixed* or a *staged beamformer*.

A simplified block diagram of a hybrid beamforming receiver is shown in Figure 7. The shown receiver has four antennas and two receive chains. Even though hybrid beamformers typically employ phased arrays, element-wise amplifiers are shown for

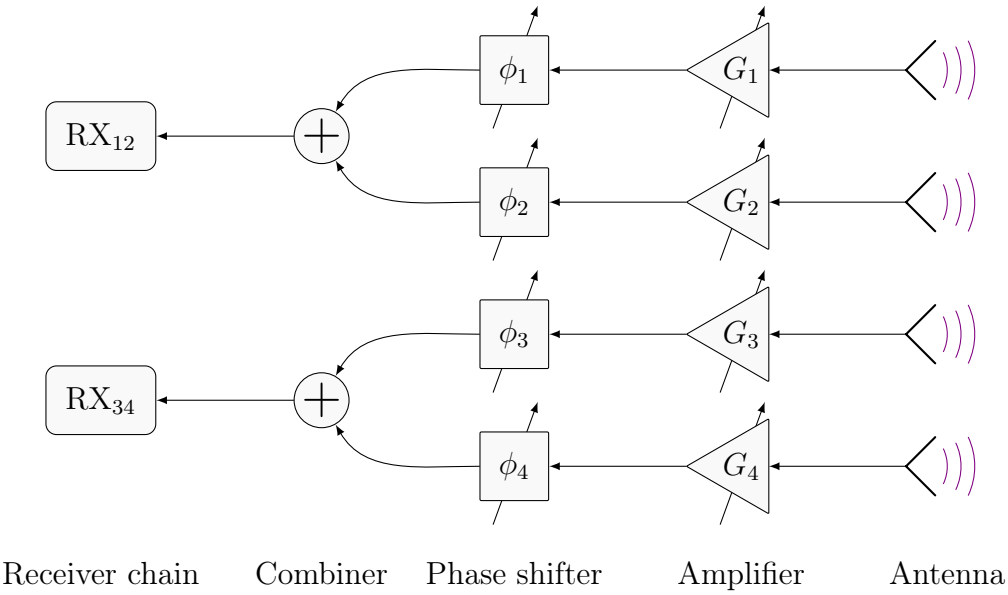


Figure 7: Hybrid beamforming receiver with four antennas and two distinct side-by-side subarrays with dedicated receiver chains.

completeness. While the configuration illustrated in Figure 7 employs side-by-side subarrays, interleaved or mixed geometry is also possible [77], [79], [92]. Furthermore, the subarrays may be distinct or indistinct. In the *distinct* case, each antenna element is connected to exactly one transceiver like shown in Figure 7. The *indistinct* configuration is similar to conventional MIMO systems where each antenna element may be connected to multiple transceivers [76], [78], [80]. In the latter configuration, there is an analog weight for each element–transceiver pair.

Using multiple transceiver chains improves the analog beamforming platform and achieves significant performance improvements [28], [37], [67], [76]–[78], [80], [92]. DSP may also be used to mitigate the effect of some of the drawbacks inherent in analog phase shifting systems. For instance, the shifting may be increasing range using digital delays. Furthermore, HBFs support SU-MIMO, P2MP functionality or even both, i.e., MU-MIMO, by using multiple independent streams simultaneously. These improvements are obtained at an acceptable cost since the number of transceiver chains remains limited.

2.3 Other mechanisms

Next, we study beam steering mechanisms that do not involve complex feed control. Three different approaches are investigated, and all of them employ array-like structures. They are, however, not used in the same fashion as those studied before. This time, only one feed is active at any moment in time. Two of the approaches use an additional focusing element while the third uses a traditional array fed through a fixed feed network.

When using one radiating element with limited directivity, the radio waves disperse into multiple directions resulting in low directivity. To avoid this, additional focusing structures are often used. The purpose of such structures is to convert a spherical wave into a directed planar wave [54]–[56], [62], [93], [94]. Referred to as *collimation*, this process is based either on reflection or on refraction when implemented using a reflector or a dielectric lens, respectively. In metallic lenses, refraction is emulated using a constrained, i.e., guiding, reflection. Hence metallic lenses are also referred to as constrained lenses.

One way to make collimating antennas steerable is to move the feed with respect to the focusing element [54]–[56], [62]. If the antenna is fed at an offset from its focal point, the beam is not aligned with focal axis anymore but turns. One can move the feed physically, or switch electronically between multiple fixed feeds placed within the focal area and its immediate vicinity.

This work considers only the latter option, referred to as *beam switching* [55]–[58], [93]. Using multiple feeds is preferred due to increased reliability and faster steering compared to mechanical means. As the feeds are internally organized into an array, one could – in principle – perform beamforming or at least use multiple independent beams. However, this is not done in practice as only one feed is typically active at any given time. This can be explained with smaller illumination area and more complex, i.e., lossy, feed network, both of which result in decreased antenna gain.

Comprising of a simple switched feed network, beam switching is easy to implement and operate [55]–[58]. While it is similar to codebooked beamforming, there is no need to control element-wise excitation; it is only necessary to select the used feed. On the other hand, there are some drawbacks. Collimating structures are often big and bulky, and have limited tuning range. Moreover, the steering is stepwise with as many options available as there are feeds. Further, the beamwidth cannot be tuned but is fixed. Regardless, beam switching is a commercially attractive solution due to its maturity and cost-effectiveness [95].

An example of a beam switched antenna is shown in Figure 8. The antenna in question is an *integrated lens antenna* (ILA) employing an internal multi-feed array [93], [94], [96]. Attached directly to the lens, this focal array is often implemented using microstrip patch antennas. Placed directly front of the array, the lens consists of an *extension* and an elliptical or hemispherical *collimator*.

Point-to-point radio links based on this technology are already commercially available. The Mayak-1000 system by RadioGigabit [95] incorporates an ILA, 60 GHz FDD transceiver and a modem to provide throughput of 1.2 GHz. The ILA has a diameter of 80 mm, a gain of 32.5 dBi and half-power beamwidth of 3.6 °.

Another approach to making collimating antennas steerable is to use a tunable collimating structure [55], [56], [58]. For this purpose, an *electronic mirror*, i.e., a reflectarray, or an *electronic lens* may be used. Both are based on the same operating principle of simulating a phased antenna array using a wireless feed network with just

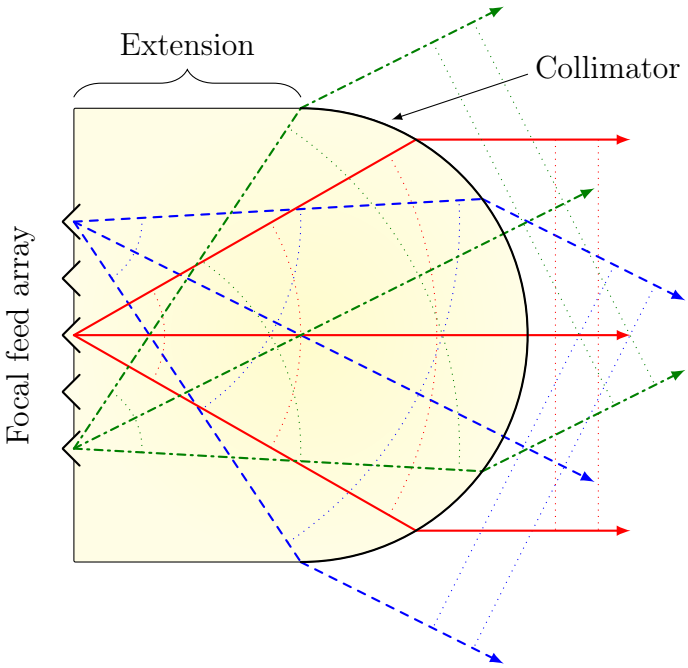


Figure 8: Beam switching integrated lens antenna with five-element feed array.

one input port, referred to as *space feeding*. The element-wise phasing traditionally done in the feed network is now carried out in the focusing element instead.

This idea is illustrated in Figure 9, depicting a space-fed five-element electronic lens antenna [55], [56], [58]. A type of *constrained lens antenna*, such antennas are also referred to as *active lens antennas*. A single radiator is directly connected to the transceiver without any extra circuitry. When used as a transmitter, this low-gain radiator creates a spherical wave captured by the inward lens antennas. As these antennas are placed on a focal arc, the lens shown in Figure 9 is also known as a *bootlace lens*. The captured signal is fed to the corresponding out-facing antenna through a phase shifter or a delay line to create a directed wave. Due to bidirectional operation and limited available space, typically only phase shifters are used.

Being in essence a phased array with a single transceiver, tunable collimation provides same functionality [55], [56], [58]. For instance, steering range is often greater compared to the case where regular collimator is used. In addition to these, complex feed networks are avoided, and the lens is – in principle – interchangeable. On the downside, advanced beam control features, such as variable beamwidth and null placement, are not supported. Since there is only a single transceiver, the system is incapable of multistream operation. However, dual-polarization extension is possible in theory, but this would further complicate the already complex and lossy design.

A phased antenna array may also be created using a fixed feed network [55]–[58], [60], [62]. In this case, the beam is steered similarly as an ILA – by switching between multiple feeds. The feed network has multiple input ports and the element-wise

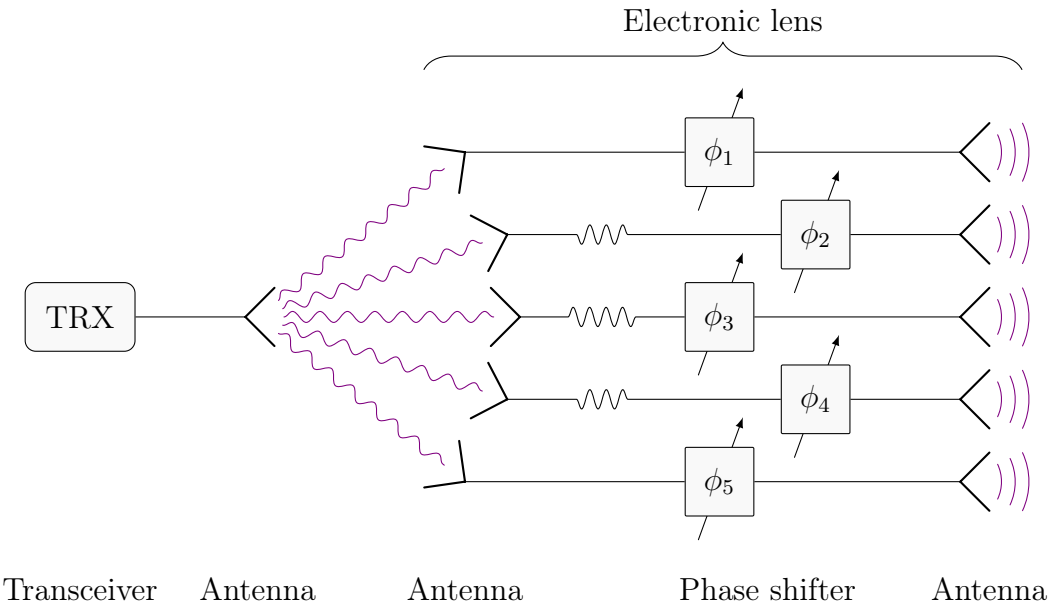


Figure 9: A space-fed five-element electronic bootlace lens.

phasing is determined by the used input port. Common features to all such structures include step-wise steering, a reasonable steering range, fixed beamwidth and support for multiple independent beams. In practice, these structures are often limited to linear arrays.

Two such implementations are showcased here, starting with the *Butler matrix* [55]–[58], [60], [62]. The Butler matrix is a $N \times N$ network ($N \in \mathbb{N}^+$), capable of generating N orthogonal, symmetrically distributed beams. An example of a 4×4 Butler matrix is shown in Figure 10, constructed using fixed $\pm\pi/4$ phase shifters and 3 dB couplers also referred to as hybrids. Both $90\,^\circ$ and $180\,^\circ$ hybrids may be used to perform the spatial fast Fourier transform (FFT). While two-dimensional spatial FFT is indeed possible, the matrix is quite complex and limited in bandwidth even in one dimension.

The other structure showcased here is referred to as the *Rotman* or *Rotman-Turner lens* [55], [56], [58], [60], [97]. It has been further developed into an *Archer lens* with some additional features. Both lenses are two-dimensional *trifocal bootlace lenses* typically implemented on a microstrip patch. It comprises a flat wave guiding medium that is symmetrical about its broadside axis, and four sets of ports on three focal arcs.

In a $N \times M \times L$ Rotman or Archer lens ($N, M, L \in \mathbb{N}^+$), there are N beam ports for inputs, two identical sets of dummy ports with M ports each, and L array ports for outputs [55], [56], [58], [60], [97]. While three-dimensional multifocal bootlace lenses are possible in theory, practical implementations are limited to linear arrays. It is, however, possible to stack multiple lenses to implement two-dimensional scanning.

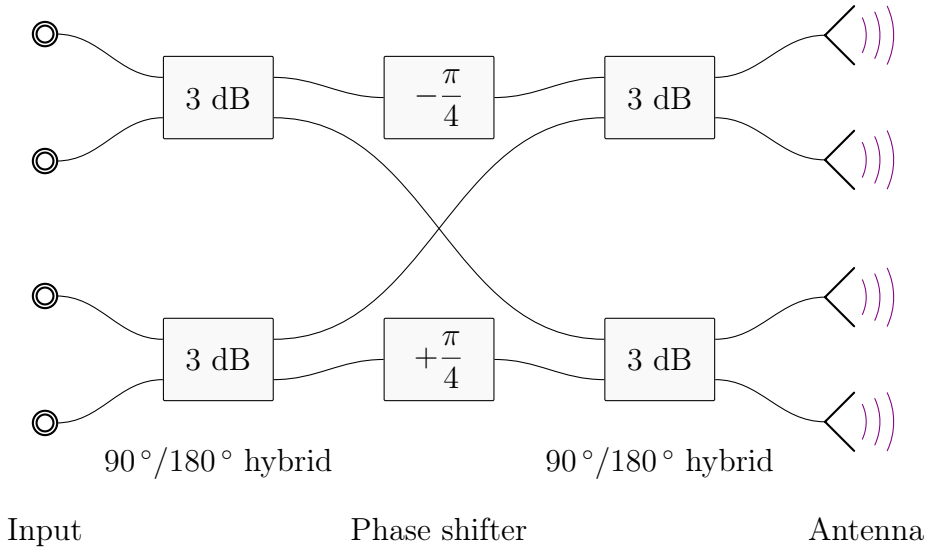


Figure 10: A 4×4 Butler matrix.

An example of a $9 \times 8 \times 8$ Rotman lens is pictured in Figure 11 [98]. Designed to operate at 60 GHz, the dimensions of this lens are 38×43 mm^2 (width and height, respectively). The nine beam ports on the left are connected to the transceiver through a switching network, and the eight array ports on the right are connected to individual antennas. The broadside axis is thus the horizontal axis, and the 2×8 dummy ports on the top and bottom are terminated using matched terminators. Switching between the beam ports determines the beam direction.

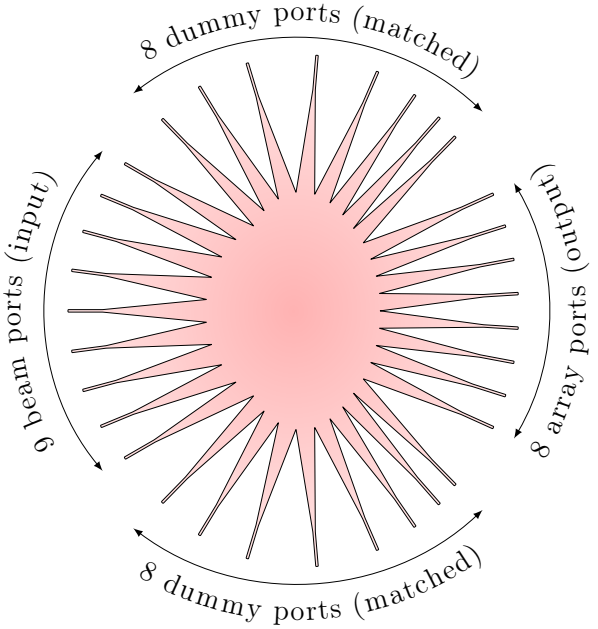


Figure 11: A $9 \times 8 \times 8$ Rotman lens [98].

3 Propagation considerations

This section concerns the radio channel and the propagation characteristics of the target application. This discussion is of utmost importance since the channel forms the basis for the entire network; it dictates a set of constraints the design needs to adhere to [87], [99], [100]. This involves not only adverse characteristics and challenges to overcome for required availability but also features to exploit for improved performance.

This section covers five main topics. The first subsection studies the free-space characteristics, discussing effects that are always present and thus concern all radio links. While this assumes an ideal environment, the actual deployment environment changes the situation considerably. This thesis considers an urban street canyon environment described in Section 3.2. Section 3.3 takes a more mathematical approach and considers the end-to-end transfer function that includes the channel effects. The millimeter-wave propagation effects and channel models that capture this behavior are discussed in Sections 3.4 and Section 3.5, respectively.

3.1 Free-space

Next, we consider the propagation characteristics of millimeter waves in free-space – an infinitely large environment that is empty except for two antennas at the two ends of the link. While such an ideal environment is purely theoretical and not found in practice, it can be used to provide some key information. This includes the free-space path loss due to expanding spherical waves and statistical phenomena due to atmospheric effects such as attenuation caused by both atmospheric gases and rainfall.

Fixed point-to-point (P2P) microwave radio links are typically designed such that the low-order Fresnel clearance zones are free of any objects [86], [87], [101]. In this cases, line-of-sight (LOS) behavior dominates received signal characteristics, and the situation can be approximated as free-space. A number of Fresnel zones are defined such that a reflection from the surface of the N^{th} causes a $180N$ ($N \in \mathbb{N}^+$) degree phase delay. The radius of N^{th} Fresnel ellipsoid F_N at any point in the LOS path may be found as

$$F_N = \sqrt{\frac{N\lambda r_1 r_2}{r}}, \tag{6}$$

where λ is the wavelength, r_1 and r_2 are the LOS distances from the point of interest to the two antennas and $r = r_1 + r_2$ is the total link length [86], [87], [101].

Consider an example with a 100 m terrestrial link at 60 GHz. In this case, the maximum radius of the first Fresnel ellipsoid is $F_{1,\,\text{max}} = 35$ cm. This is easily satisfied given the small hop distance and wavelength. As a result, LOS behavior can be expected to dominate as long as there is a clear LOS connection. However, this might not always be the case given the targeted use cases, and certainly not the whole picture as will be explained in the following subsections.

Assuming a LOS-dominant scenario discussed above, the received power may be estimated as follows [53], [85]. The transmitted radio waves expand in a spherical fashion, and the magnitude of the electric field is inversely proportional to the distance. The energy flux density may be found using *Poynting vector* $\vec{S} = \vec{E} \times \vec{H}$ where $\vec{E}$ and $\vec{H}$ are the electric and magnetic field vectors, respectively, and $\times$ is the vector cross product. Interested in the magnitude $|\vec{S}|$, it may be found using transmit power delivered to the antenna P_{T}, transmit antenna gain G_{T}, and distance r as

$$|\vec{S}| = G_{\mathrm{T}} \frac{P_{\mathrm{T}}}{4\pi r^2}. \tag{7}$$

The received power depends on how much of this energy flux the receive antenna captures. The received power available from the antenna P_{R} may be found using the effective antenna area A_{eff} as $P_{\mathrm{R}} = A_{\mathrm{eff}}|\vec{S}|$. The effective antenna area is defined as

$$A_{\mathrm{eff}} = G_{\mathrm{R}} \frac{\lambda^2}{4\pi}, \tag{8}$$

where G_{R} is the receive antenna gain toward the transmitter and λ the wavelength in used transmit medium. The received power is thus

$$P_{\mathrm{R}} = G_{\mathrm{R}} G_{\mathrm{T}} \left(\frac{\lambda}{4\pi r} \right)^2 P_{\mathrm{T}}, \tag{9}$$

commonly referred to as the *Friis transmission equation* [53]–[55], [85].

Equation (9) is highly simplified and does not take atmospheric effects into account [72]. Moreover, it does not consider antenna alignment; the gain variation and polarization mismatch [54], [55]. The equation may, however, be extended to include these features. Using logarithmic quantities, i.e., decibels (dB) and decibel-milliwatts (dBm), the *extended Friis formula* may be written as

$$\begin{aligned} P_{\mathrm{R,\,dBm}} = {} & P_{\mathrm{T,\,dBm}} + G_{\mathrm{R,\,dB}}(\vartheta_{\mathrm{R}}, \varphi_{\mathrm{R}}) + G_{\mathrm{T,\,dB}}(\vartheta_{\mathrm{T}}, \varphi_{\mathrm{T}}) \\ & - 20 \log_{10} \left(\frac{4\pi r}{\lambda} \right) \mathrm{dB} - L_{\mathrm{pol,\,dB}} - r L_{\mathrm{atm,\,dB}} - r L_{\mathrm{rain,\,dB}}. \end{aligned} \tag{10}$$

The $\vartheta_{\mathrm{R}}, \vartheta_{\mathrm{T}}$ and $\varphi_{\mathrm{R}}, \varphi_{\mathrm{T}}$ are the elevation and azimuth angles at both ends of the link, respectively. $L_{\mathrm{pol,\,dB}}$ is the loss due to polarization mismatch between the antennas at each end. Finally, $L_{\mathrm{atm,\,dB}}$ and $L_{\mathrm{rain,\,dB}}$ are the attenuations due to atmospheric gases and rainfall, given in decibels per unit length, respectively.

While, Equation (10) is a less crude simplification of the situation, it does not consider all phenomena affecting received signal quality [86]. These phenomena are out of the scope of this thesis but the interested reader may resort to Ghasemi's book on radio wave propagation [86] or Recommendation P.530-16 [101] on link planning and accurate prediction modeling by the Radiocommunication Sector of the International Telecommunication Union (ITU-R).

Equation (10) can regardless be used to estimate received power and forms the basis of link budget calculations [53], [85], [87], [101]. The equation, however, does

not actually highlight when the signal is successfully recovered. The link is possible only if signal to noise and interference ratio (SINR) is high enough [53], [85], [87], [101]. Thus, noise and interference originating from the environment and within the receiver should be taken into account.

The terms in Equation (10) may be divided into two parts shown here on their own lines. On the first line, we have the antenna gains and transmit power, all of which can be tuned within certain limits through beam steering and gain control. As long as the unwanted signals remain uncorrelated to the transmit signal, these terms have an increasing effect on the SINR [85]. As a result, antenna gain and transmit power may be used to compensate for the attenuation terms on the second line. Determined by the deployment setting and environment, these terms are fixed and have a negative effect on the SINR.

There are four of these loss terms, the first and most significant of which is the free-space path loss (FSPL). The FSPL corresponds to the loss due to expanding spherical wave; the power is distributed over a larger area [53]–[55], [85]. It also includes the antenna capture area, meaning the FSPL is proportional to both frequency f and distance r squared since $\lambda = c/f$:

$$FSPL \propto (fr)^2. \tag{11}$$

Considering the previous example with a 100 m terrestrial link at 60 GHz, the FSPL equals to 108 dB.

In addition to FSPL, there is a term for polarization mismatch loss L_{pol}. Assuming linear polarization, it is defined as a dot product $(\cdot)$ of the polarization unit vectors $\vec{p}_{\text{T}}$ and $\vec{p}_{\text{R}}$ at the two ends of the link [54], [55]. It further reduces to cosine of the polarization tilt angle τ:

$$L_{\text{pol}} = (\vec{p}_{\text{T}} \cdot \vec{p}_{\text{R}})^{-2} = \cos^{-2}\tau. \tag{12}$$

In addition to direction-dependent antenna gain, L_{pol} describes the third axis of antenna rotation.

Loss due to polarization mismatch can usually be neglected in the context of fixed point-to-point links such as backhaul, as the polarizations are aligned during installation within reasonable accuracy [54], [55], [84]. For instance, a two-degree misalignment causes a negligible loss of 0.0053 dB. While fixed links can be aligned simply by using a level, polarization mismatch can cause significant losses in access links where the user may hold their device in arbitrary orientation. Thus, the effect of mismatch losses is typically mitigated with polarization diversity techniques [84].

The third term in Equation (10) is the attenuation due to atmospheric gases, namely due to air and water vapor. The attenuation experienced by the radio waves is dependent on frequency, height, temperature, air pressure and humidity. It can be modeled using ITU-R Recommendation P.676-10 [102] specified for frequencies above 1 GHz. This model is shown in Figure 12 with graphs for different causes. The figure is plotted up to 99 GHz at sea level in standard barometric pressure

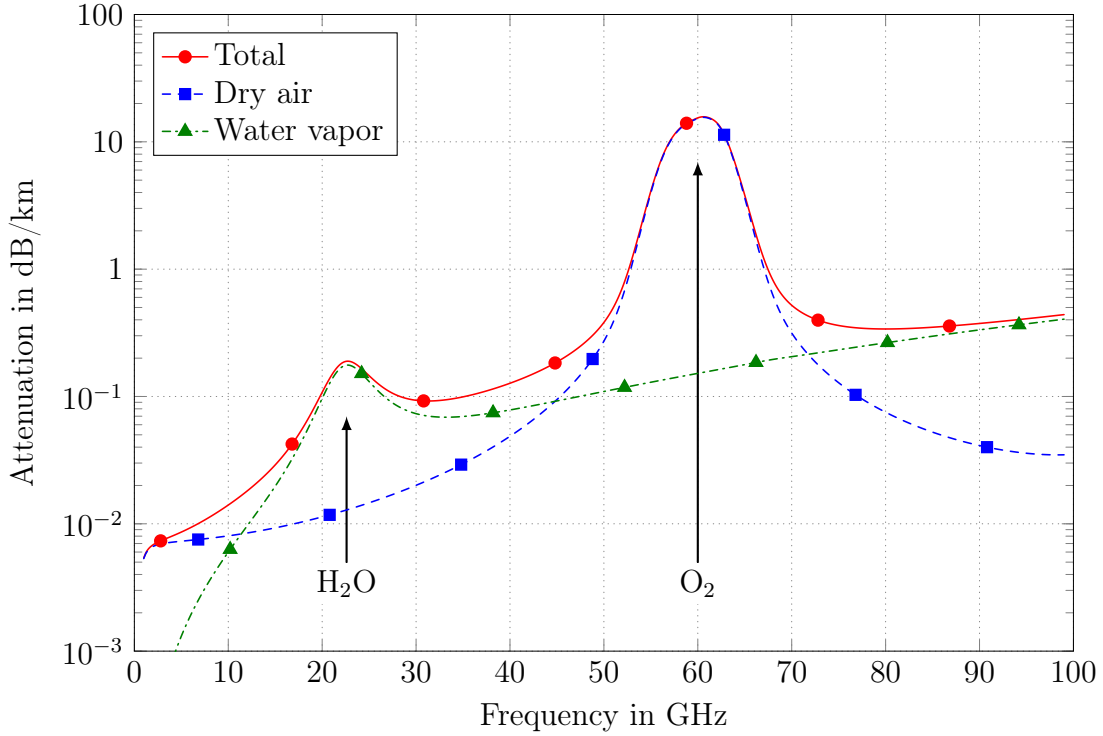


Figure 12: Attenuation due to atmospheric gases as a function of frequency [102].

1013.25 hPa, and the air temperature and water vapor density are set to 15 °C and 7.5 g/m^3, respectively. Around 60 GHz, the attenuation peaks at 15.5 dB/km due to oxygen (O_2) absorption.

The fourth and last attenuation term in Equation (10) is due to rainfall and fog. The term is naturally dependent on the precipitation rate and fog thickness, i.e., water density, which varies over time. Aiming to meet a certain availability goal, the corresponding rate is used in link budget calculations [85], [87], [101]. Based on years of medium-range weather forecasts, the expected or average rates at given location may be determined using the ITU-R Recommendation P.837-6 [103].

However in this thesis, an older, coarser but also more illustrative model is used. The model in question is the ITU-R PN.837-1 [104] where the world is divided into 15 rain zones or regions $\{\mathrm{A}, \mathrm{B}, \ldots, \mathrm{Q}\} \setminus \{\mathrm{I}, \mathrm{O}\}$. These zones are highlighted in Figure 13, and the corresponding expected precipitation rates for different probability levels are given in Table 1. For instance, Southern Finland belongs to Region E where it can be expected to rain more than 22 mm/h less than 0.01 % of the time.

Knowing the precipitation rate ρ, the attenuation per unit length $L_{\text{rain, dB}}$ may be found using the power-law

$$L_{\text{rain, dB}} = \kappa \rho^{\alpha}, \tag{13}$$

where κ and α are coefficients dependent on frequency and polarization. These coefficients can be found using the ITU-R Recommendation P.838-3 [105] for frequencies above 1 GHz. The attenuation due to fog is defined similarly in ITU-R

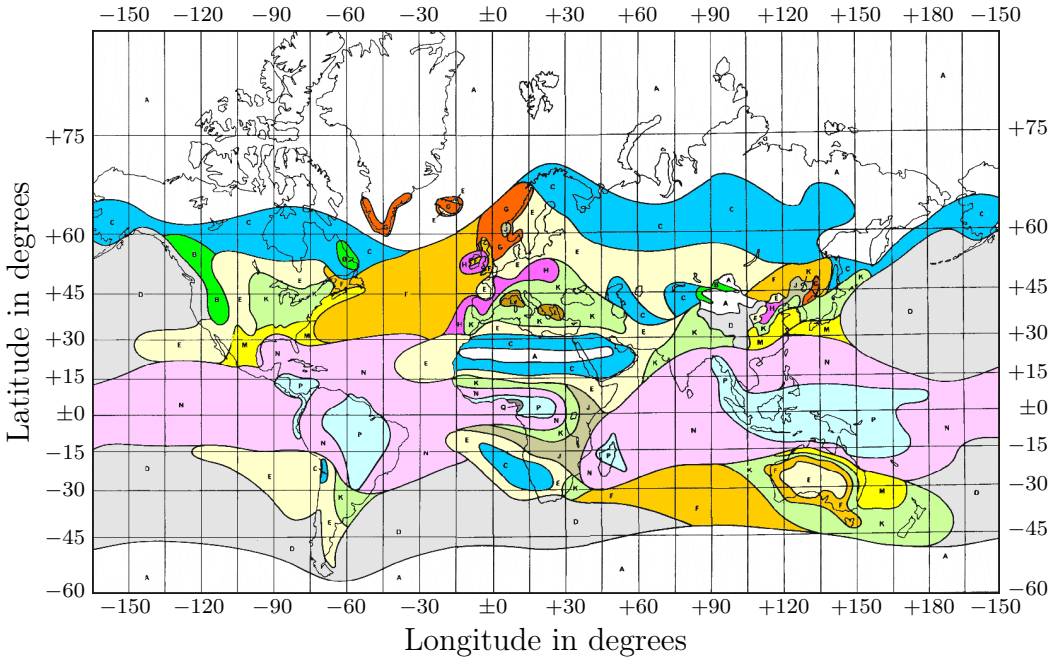


Figure 13: ITU-R PN.837-1 rain zones with rates given in Table 1 [104].

Table 1: Precipitation rates in different ITU-R PN.837-1 rain zones in mm/h [104].

%	A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
1.0	<0.1	0.5	0.7	2.1	0.6	1.7	3	2	8	1.5	2	4	5	12	24
0.3	0.8	2	2.8	4.5	2.4	4.5	7	4	13	4.2	7	11	15	34	49
0.1	2	3	5	8	6	8	12	10	20	12	15	22	35	65	72
0.03	5	6	9	13	12	15	20	18	28	23	33	40	65	105	96
0.01	8	12	15	19	22	28	30	32	35	42	60	63	95	145	115
0.003	14	21	26	29	41	54	45	55	45	70	105	95	140	200	142
0.001	22	32	42	42	70	78	65	83	55	100	150	120	180	250	170

Recommendation P.840-6 [106]. Unlike rain loss, attenuation due to fog is directly proportional to water density.

Attenuation at a number of different precipitation rates and fog thicknesses are plotted in Figure 14 up to 99 GHz assuming vertical polarization. At 60 GHz, rain and fog cause attenuation in excess of 10 dB/km and 1 dB/km when the precipitation rate and water density are above 30 mm/h and 0.5 g/m^3, respectively. More susceptible to rain, horizontally polarized signals would experience greater attenuation, leading to imbalance in dual-polarized systems with long hops.

The ITU-R P.838-3 rain model [105] functions by approximating the effective permittivity of the transport medium, i.e., air, using average water densities derived from the average precipitation rate over some hours. Even though based on extensive empirical studies, it does not consider different types of rain nor its inhomogeneity.

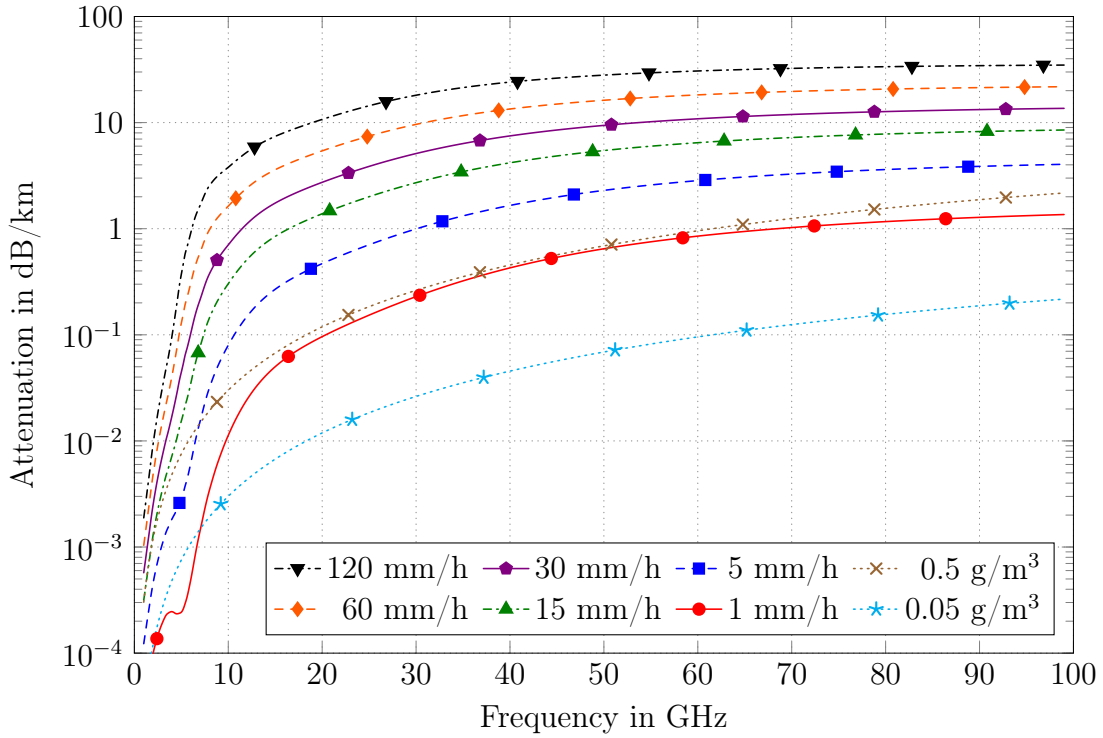


Figure 14: Attenuation due to rainfall and fog as a function of frequency at different precipitation rates (mm/h) and water densities (g/m^3), respectively [105], [106].

Same total rainfall may be a result of various specific rain profiles, especially in distinct geographical locations with their own climate but also within same area [107].

With increasing interest toward millimeter-wave radio links and more significant effect on link quality, researchers are questioning validity of averaging approaches [108]. The models are mainly intended for characterizing macro-level behavior in fixed long range communications over large distances and over long durations of time [107]. As such, they do not capture highly localized effects that are significant for relatively short hops.

Moreover, ITU-R P.838-3 attenuation model does not consider the effect of individual raindrops even though they are of comparable size when compared to the wavelength [108]. As a result, their size, distribution, temperature and shape, and even wind affects the resulting attenuation and other effects, such as polarization mixing, not captured by the models. Furthermore, the models do no take other precipitation effects such as hail or snow, or icing into account [101]

This subsection concludes with a comparison of free-space propagation characteristics between millimeter waves and conventional frequencies. The differences are summarized in Table 2, and in Figure 15. Nominal values are shown in the table while the figure highlights changes in availability requirements using ITU-R recommendations P.676-10 [102], PN.837-1 [104] and P.838-3 [105]. A 1 km outdoor link is investigated at sea level and at standard atmospheric pressure of 1013.25 hPa.

The temperature and water vapor density are 15 °C and 7.5 g/m^3, respectively. The radio signals are assumed to be vertically polarized.

Table 2 shows the wavelength in air λ, the maximum radius of the first Fresnel ellipsoid F_1, free-space path loss *FSPL*, attenuation due to atmospheric gases L_{atm} and rainfall L_{rain} assuming 22 mm/h precipitation. In addition to these parameters, antenna gain corresponding to 0.1 m^2 effective area is given. The antenna gain corresponding to a certain effective antenna area is found by solving Equation (8) for G.

The values highlight the challenging radio environment at millimeter waves compared to conventional frequencies. Considering the unlicensed bands at 2.4 GHz and 60 GHz, the total attenuation is 52 dB higher at millimeter waves. Fortunately, antenna gain is simultaneously increased by 28 dB which compensates the increase in FSPL since both are proportional to the frequency squared. Considering the entire link, the antenna gain increase needs to be doubled to account for both ends. As a result, the received signal is 4 dB stronger at 60 GHz.

However, the antenna gain values shown in Table 2 are very optimistic and thus highly unlikely to be achieved in practice. The effective area is an idealization, dependent on both the physical size of the antenna and its efficiency among other factors [54]–[56], [62]. Even if the effective area would match the physical size, 0.1 m^2 corresponds to 333×333 mm^2 which is simply too large given the form factor

Table 2: Comparison of free-space propagation characteristics assuming a 1 km terrestrial link.

Parameter	**Symbol**	**Unit**	**1 GHz**	**2.4 GHz**	**5 GHz**
Wavelength in air	λ	mm	300	125	60.0
Clearance zone	F_1	m	8.66	5.59	3.87
Free-space path loss	*FSPL*	dB	92.4	100	106
Atmospheric loss	L_{atm}	dB	0.00538	0.00712	0.00861
Loss due to rain	L_{rain}	dB	0.000438	0.00298	0.0276
Antenna gain	G	dBi	11.5	19.1	25.4

Parameter	**Symbol**	**Unit**	**60 GHz**	**75 GHz**	**85 GHz**
Wavelength in air	λ	mm	5.00	4.00	3.53
Clearance zone	F_1	m	1.12	1.00	0.939
Free-space path loss	*FSPL*	dB	128	130	131
Atmospheric loss	L_{atm}	dB	15.4	0.361	0.350
Loss due to rain	L_{rain}	dB	8.61	9.91	10.5
Antenna gain	G	dBi	47.0	49.0	50.0

target. Furthermore, such high gain leads to extremely narrow beams, and small misalignment reduces gain considerably.

Beam steering capability usually decreases the efficiency which in turn reduces gain due to losses in extra components, e.g., insertion loss of switches [53], [84]. Also, increasing beamwidth decreases the antenna directivity and gain [54]–[56], [62]. For instance, the previously discussed Mayak-1000 system [95] by RadioGigabit incorporates an ILA with 32 dBi gain at 60 GHz, suggesting shorter hops.

Moreover, there are requirements regarding antenna gain and maximum radiated power [85], [87] among other limitations such as bandwidth. These limits are specified by local authorities for each frequency band. Equivalent isotropically radiated power (EIRP) is one of the typical criteria as it combines both transmit antenna gain and transmit power. It is defined as a product of the two [53], [86]:

$$EIRP = G_{\mathrm{T}} P_{\mathrm{T}}. \tag{14}$$

For instance in the USA, the Federal Communications Commission (FCC) has set the maximum average and peak EIRPs in the unlicensed 57 ... 64 GHz V-band to +40 dBm and +43 dBm (10 W and 20 W), respectively [109]. Similar limits are in force for both V- and E-bands all over the world [75].

Figure 15 shows the statistical attenuation as a function of probability at different frequencies assuming previously used conditions. The losses due to atmospheric gases are fixed and the rainfall statistics correspond to Southern Finland, i.e., ITU-R

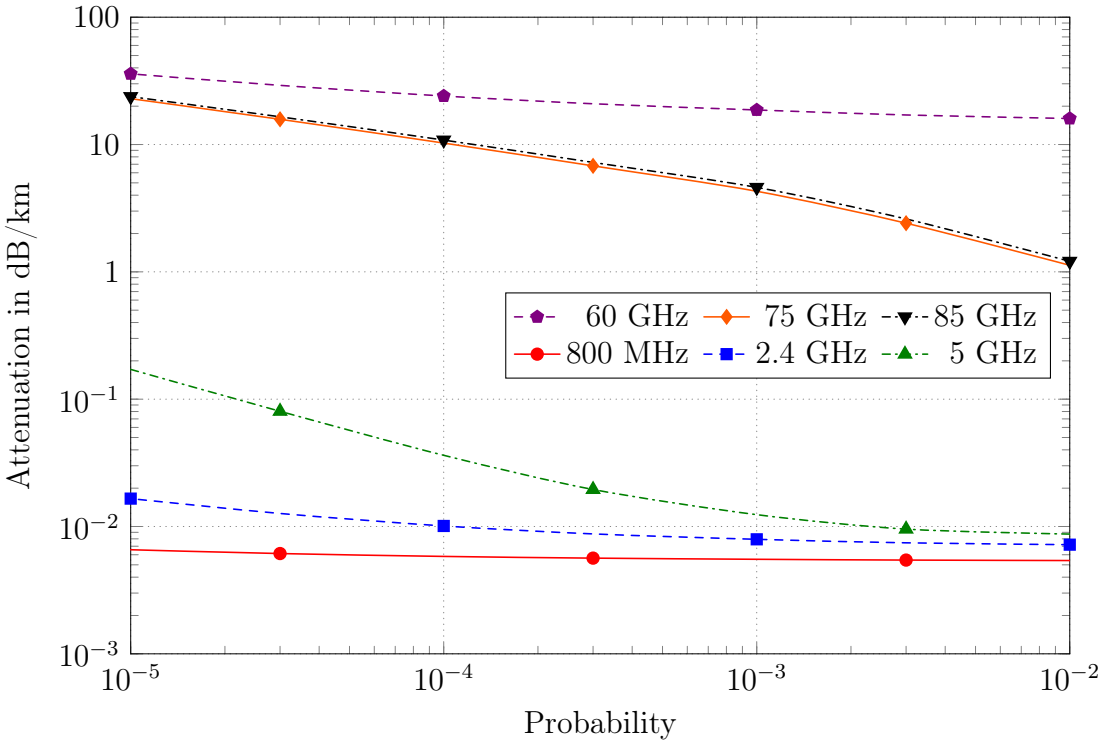


Figure 15: Attenuation as a function of probability at different frequencies [105].

P-837.1 Zone E [104] given in Table 1. The graphs clearly highlight the deleterious nature of rainfall at millimeter waves. At low frequencies, the effects of rain and fog are negligible [101]. At mmW frequencies, however, the attenuation is more than two orders of magnitude larger; exceeding 20 dB/km at 70 mm/h precipitation corresponding to 0.001 % probability.

3.2 Street canyon environment

The previous subsection considered a theoretical deployment scenario; a void empty of all objects except for the two antennas. Such a theoretical study is enough to capture the fundamental nature of radio wave propagation. Sadly this is not enough in practice because the effect of obstacles within or near the Fresnel zones [101] must also be considered. New propagation methods are introduced when radio waves reflect, diffract and scatter from or penetrate through various objects in the environment as shown in Figure 16 [46], [84], [86].

There are two consequences [84], [86], [87]. First, it is possible for the radio waves to reach the receiving end behind objects, increasing coverage. Second, the waves may reach the receiver through multiple different propagation paths, referred to as *multipath propagation*. Signals in these different paths experience different characteristics such as attenuation, delay and rotated or otherwise mixed polarization. Furthermore, when these signals from different paths are received simultaneously, received signal strength varies based on the receiver location [84], [86], [87].

This phenomenon is referred to as *fading* which can be classified into slow and fast fading. If the receiving endpoint is shadowed by large objects, the scenario is referred to as *slow fading* or *shadowing* and the received power distribution is *log-normal*. On the other hand, if there are multiple propagation paths between the TX and RX, the received power has multiple local minima and maxima, fluctuating several decibels over short distances. Referred to as *fast fading* or *multipath fading*, the channel is considered *Rice* or *Rayleigh distributed* whether there is a dominant LOS path or not.

Furthermore, the fading loss may be dependent on frequency or frequency-invariant, referred to as *frequency-selective* or *flat fading*, respectively [84], [86], [87]. Moreover, signals from different paths with long relative delays but comparable powers may overlap and cause intersymbol interference (ISI) [84], [86], [87]. Both ISI and frequency-selective fading have extremely deleterious effect on telecommunication systems. While frequency-selective fading shapes the received signal spectrum and its waveform, ISI can only be reduced by limiting the bandwidth. To reduce the impact of this unwanted behavior, wideband systems typically distribute signal to multiple carriers with smaller bandwidth [84], [86]–[88].

Multipath effects and fluctuations in received power resulting from fading are a major engineering challenge with mobile users who expect undisrupted service. This power ripple is reduced with diversity techniques based on signal replication: the same information is transferred more than once but in different conditions [84],

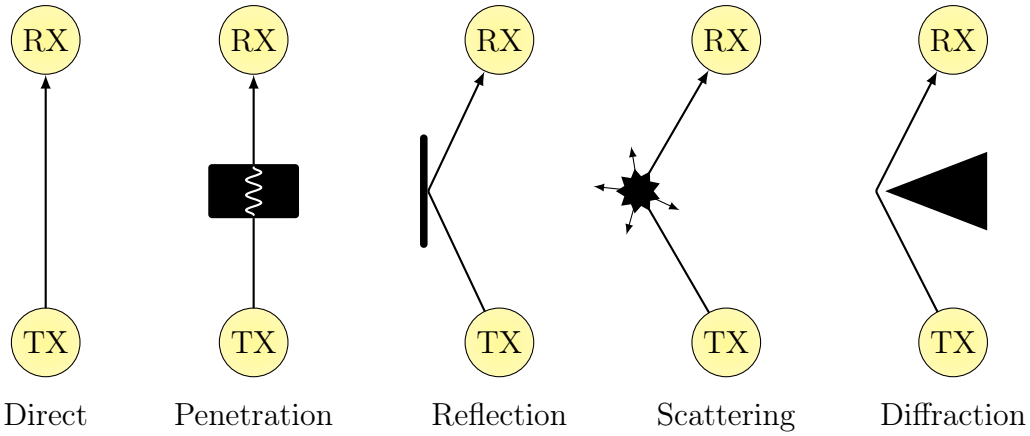


Figure 16: Different propagation mechanisms.

[86]–[89]. Even small changes in the propagation setting may be enough to avoid destructive interference due to fast fading. Three different diversity techniques are typically recognized: temporal, spatial and polarization diversity [84], [86]–[89]. These techniques involve replicating the signal in time, using multiple antennas placed apart or in different orientation, respectively.

While aforementioned power fluctuations and ISI are certainly unwanted behavior, not all multipath effects are detrimental. Instead, we can benefit from channels rich in fading by using multiple antennas and transceiver chains at both ends of the link [84], [86]–[90]. If the different propagation paths between the pairs of antennas are uncorrelated, they form multiple subchannels – a property referred to as *high spatial degree of freedom* [48], [84], [110]. These paths maybe used to transmit several independent streams in parallel, increasing throughput to one or more users. This space-time processing is referred to as spatial multiplexing, or more commonly as single-user or multiuser multiple-input-multiple-output (MIMO), SU-MIMO or MU-MIMO for short, respectively [84], [86]–[90].

Since the objects such as the terrain, vegetation, people and man-made structures are unique to a given location, also the propagation environment is unique. As a result, one cannot refer to an all-comprehensive model but must consider the target environment. While all sites are in fact unique, there are similarities between the sites. For this reason, all mobile communications standards involve a set of specified use cases with their own focus, goals and requirements [8], [9].

While next-generation mobile network standardizing is far from complete, the *Mobile and wireless communications Enablers for the Twenty-twenty Information Society* (METIS) project[2] [19] is considered the most significant contribution toward the goal. Within the project, a dozen test cases are defined [8], [9]. Targeting for different deployment scenarios, the test cases consider different terrain, structures, number of users along with their behavior and needs. They range from typical

[2]For more information on the project, please refer to `https://www.metis2020.com/`.

indoor office environment to stadiums and shopping centers, and from different traffic settings to natural disasters and other emergencies.

This thesis considers the setting in *METIS Test Case 2: Dense Urban Information Society* [8], [9] – highly-populated metropolitan areas most likely to be targeted for the very first 5G installations. The test case is further divided into three bidirectional propagation scenarios, namely indoor-to-indoor (I2I), outdoor-to-outdoor (O2O) and outdoor-to-indoor (O2I). These are further divided into three categories based on the types of devices involved: macro-to-user, micro-to-user and ad-hoc user-to-user. In total, eight scenarios are defined as macro BSs are installed on roof tops and not indoors.

The street canyon environment comprises streets, openings and high-rise buildings forming canyon floor and walls. Targeting for increased capacity, multiple mmW access points are mounted on building facades, lamp posts and other street furniture some meters off the ground. Within METIS Test Case 2, the environment is modeled using a three-dimensional *Madrid grid* model [111]. Named after the capital of Spain, Madrid grid model takes after European metropolitans. Unlike previously used models, such as the *Manhattan grid*, the model involves more lifelike heterogeneous building layout and a lively, dynamic world with realistic user behavior.

A three-dimensional render of the Madrid grid is shown in Figure 17. A to-scale layout of the model is also shown in Figure 18. Measuring $387 \times 552\,\mathrm{m}^2$, the model includes a repeatable or tileable 4-by-4 grid. Within this grid, there are 15 buildings of two different shapes ($120 \times 120\,\mathrm{m}^2$ and $30 \times 120\,\mathrm{m}^2$) rising 8 to 15 stories high, and a park ($120 \times 120\,\mathrm{m}^2$). These are surrounded by 3 m sidewalks, and there is also

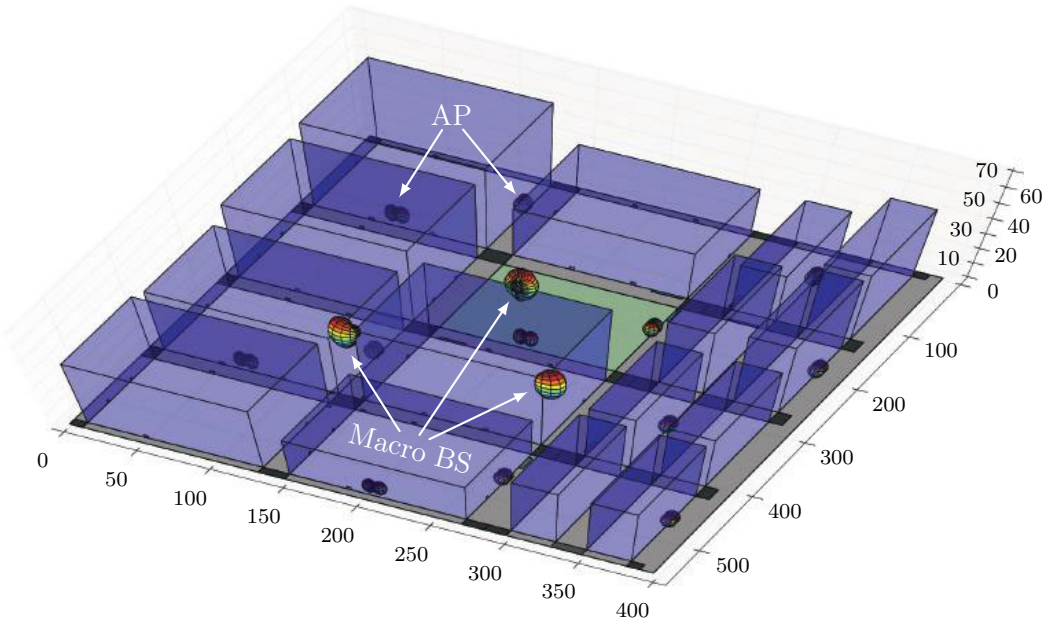


Figure 17: A three-dimensional render of the Madrid grid with dimensions in meters [112].

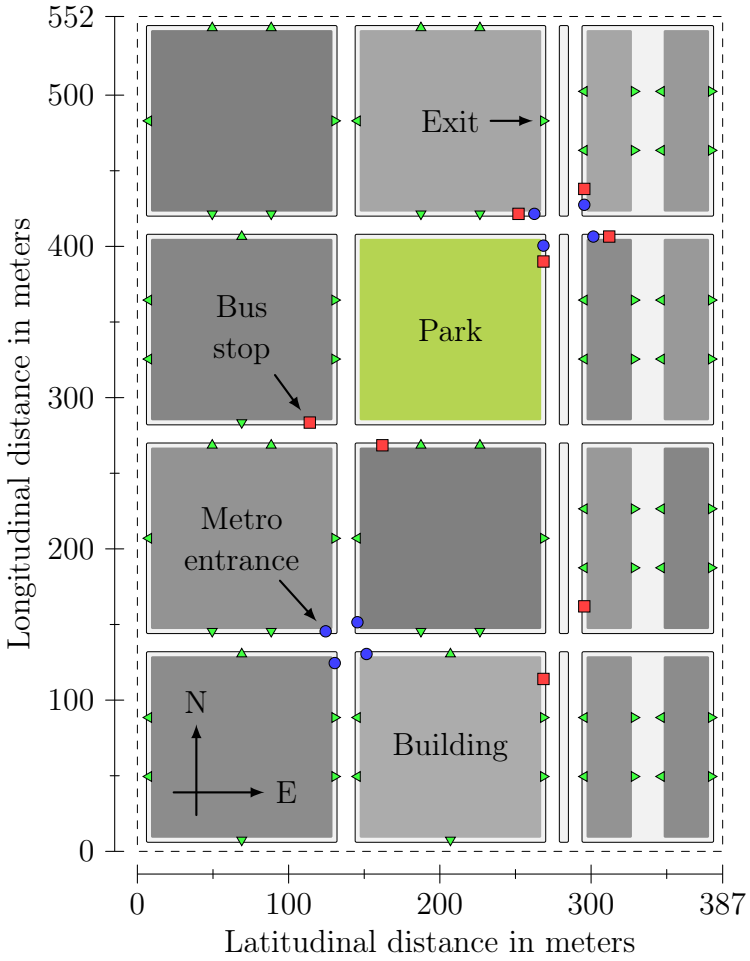


Figure 18: A two-dimensional 5000 : 1 visualization of the Madrid grid.

a 24 m thoroughfare, a 21 m promenade and regular roads 12 m wide. The model does not include any fixed installations other than the buildings.

The Madrid grid model also defines the BS and AP placement as shown in Figure 17. There are three macro base stations installed on the rooftop of the square building south of the park. They are placed on the building edge, facing north, south east and south west at a height of 52.5 m from the ground. Within the 21.3 hectares, there are 12 access points mounted on building facades 5 m above the street-level. The radio links are determined using ray tracing principles to account for realistic multipath conditions even for access links. The corresponding channel model [48] is discussed in Section 3.5.

In addition to AP and BS placement, the Madrid grid incorporates UE placement functionality. This mobility model accounts for simplified but realistic behavior of the users. There are two models, a simple model and more complex alternative, both of which include pedestrians, passenger cars and two types of public transport: bus and metro. All user and their interaction is generated using a random process with

certain constraints regarding, for instance, speed, density, traffic lights and flow of traffic.

For instance, take the complex model where users are randomly generated at building exits with the nearest public transportation stop as their destination. In Figure 18, the eight bus stops and metro entrances are highlighted using red squares (□) and blue circles (○) whereas the building exits are marked with green triangles (△). At the stops, the users wait to board the bus or the metro before continuing their journey. Typically only one grid tile is used, and moving users are folded or wrapped at the edges, continuing from the opposite side. Eventually, the users get off the bus or metro on a randomly selected stop.

The discussion above has only covered an overview of the Madrid grid model. For instance, we have not discussed different traffic scenarios such as a traffic jam. For a more complete description of the Madrid grid model, including all simulation parameters, the interested reader should refer to *METIS Deliverable D6.1 – Simulation guidelines* [111] or Pathan's book on network simulation [112].

3.3 End-to-end transfer function

The propagation of millimeter waves in a realistic environment is discussed next. The focus is on outdoor scenarios interested in radio wave behavior in a street canyon environment like described in the Madrid grid model discussed in Section 3.2. That is, how the various objects and multipath propagation change situation compared to free-space discussed in Section 3.1.

The discussion on the topic is divided into three parts. First, in this subsection, the received signal is expressed mathematically in the general case using an end-to-end transfer function. Second, some key characteristics of mmW propagation are pointed out in Section 3.4. Differences between low-frequency and mmW behavior are highlighted with an extra emphasis on topics related to beam steering. Finally, Section 3.5 showcases two different channel models intended for modeling millimeter-wave propagation in a street canyon environment.

The end-to-end transfer function between the two devices may be expressed using matrix operations [57], [59]. As the presentation is mathematically intensive, we start by declaring notation. **Upright boldface** characters to designate arrays; lowercase letters like $\mathbf{a}$ are vectors whereas capital letters like $\mathbf{A}$ are matrices. $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{H}}$ are regular and conjugate transposes of matrix $\mathbf{A}$, respectively. $\mathbf{A}\,\mathbf{B}$ and $\mathbf{A} \odot \mathbf{B}$ are regular and Hadamard, i.e., element-wise, products of matrices $\mathbf{A}$ and $\mathbf{B}$. Out of the two, Hadamard product has higher operator precedence.

This presentation considers only single-user scenarios and neglects all non-idealities inherent in the devices, such as noise, nonlinearity, crosstalk, spurious emissions or coding errors. In other words, we assume ideal devices but non-ideal environment, taking both interferers and noise generated in the environment into account. Within these constraints, the approaches taken by Alkhateeb *et al.* [76], [113] and Larew *et al.* [114], [115] may be combined to write the received signal as the matrix sum of

the wanted and unwanted signals:

$$\mathbf{X} = {\mathbf{V}_\mathrm{R}}^\mathrm{H} \odot {\mathbf{W}_\mathrm{R}}^\mathrm{H} \odot \mathbf{H} \odot \mathbf{W}_\mathrm{T} \odot \mathbf{V}_\mathrm{T}\, \mathbf{S} + {\mathbf{V}_\mathrm{R}}^\mathrm{H} \odot {\mathbf{W}_\mathrm{R}}^\mathrm{H} \odot \mathbf{I}_2\, \mathbf{N}. \tag{15}$$

Here $\mathbf{X}$ and $\mathbf{S}$ are the received and transmitted signal matrices, respectively. $\mathbf{V}$ and $\mathbf{W}$ are the baseband combiner-precoder and antenna weight matrices, respectively. Subscripts 'T' and 'R' are used to refer to the transmitter and receiver, respectively. Depending on the definition, the receiver weight matrices may be transposed but not conjugated in some context [72]. $\mathbf{H}$ is the channel matrix, $\mathbf{N}$ is the unwanted signal matrix and $\mathbf{I}_2$ is the two-by-two identity matrix used to match dimensions:

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Assuming a dual-polarized system, all aforementioned terms are matrices, not vectors. Furthermore, both weight matrices $\mathbf{V}$ and $\mathbf{W}$ are needed to simultaneously account for both analog and digital beamforming. Moreover, there are no constraint on the number of parallel data streams using the two polarizations. In short, Equation (15) is valid even for hybrid beamforming system employing MIMO processing using two polarizations.

In such systems, the signal streams $\mathbf{S}$ are first MIMO precoded to different transmitters with $\mathbf{V}_\mathrm{T}$ followed by transmitter to antenna mapping $\mathbf{W}_\mathrm{T}$. The radiated signals experience transformations by the channel $\mathbf{H}$ before they are captured by the receiving antennas. The signals are then passed to different receiving chains after weighting with ${\mathbf{W}_\mathrm{R}}^\mathrm{H}$. The received signals $\mathbf{X}$ are finally obtained after combining different receiver signals with ${\mathbf{V}_\mathrm{R}}^\mathrm{H}$. The receive function works similarly for unwanted signals represented with $\mathbf{N}$.

For further manipulation of transfer function, the matrices in Equation (15) are expanded. The signal matrices $\mathbf{X}$, $\mathbf{S}$ and $\mathbf{N}$ comprise of vectors for both vertical and horizontal polarizations. Using subscripts 'V' and 'H' to designate vertical and horizontal polarization, respectively, the signal matrices are defined as

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_\mathrm{V} \\ \mathbf{x}_\mathrm{H} \end{bmatrix}, \qquad \mathbf{S} = \begin{bmatrix} \mathbf{s}_\mathrm{V} \\ \mathbf{s}_\mathrm{H} \end{bmatrix} \quad \text{and} \qquad \mathbf{N} = \begin{bmatrix} \mathbf{n}_\mathrm{V} \\ \mathbf{n}_\mathrm{H} \end{bmatrix}.$$

$\mathbf{x}_\mathrm{V}$ and $\mathbf{x}_\mathrm{H}$ are the vertically and horizontally polarized received signal vectors. Transmitted signal and unwanted signal vectors $\mathbf{s}_\mathrm{V}$, $\mathbf{s}_\mathrm{H}$, $\mathbf{n}_\mathrm{V}$ and $\mathbf{n}_\mathrm{H}$ are defined similarly.

The signal matrices comprise of signal vectors for both polarizations, and the weight matrices $\mathbf{V}_\mathrm{R}$, $\mathbf{W}_\mathrm{R}$, $\mathbf{V}_\mathrm{T}$ and $\mathbf{W}_\mathrm{T}$ are built in a similar fashion. This time, however, the polarization-wise elements are matrices themselves. Moreover, they need to be replicated to match dimensions. They are defined as

$$\mathbf{V}_\mathrm{R} = \begin{bmatrix} \mathbf{V}_{\mathrm{R,V}} & \mathbf{V}_{\mathrm{R,H}} \\ \mathbf{V}_{\mathrm{R,V}} & \mathbf{V}_{\mathrm{R,H}} \end{bmatrix}, \qquad \mathbf{W}_\mathrm{R} = \begin{bmatrix} \mathbf{W}_{\mathrm{R,V}} & \mathbf{W}_{\mathrm{R,H}} \\ \mathbf{W}_{\mathrm{R,V}} & \mathbf{W}_{\mathrm{R,H}} \end{bmatrix},$$

$$\mathbf{V}_\mathrm{T} = \begin{bmatrix} \mathbf{V}_{\mathrm{T,V}} & \mathbf{V}_{\mathrm{T,H}} \\ \mathbf{V}_{\mathrm{T,V}} & \mathbf{V}_{\mathrm{T,H}} \end{bmatrix} \quad \text{and} \qquad \mathbf{W}_\mathrm{T} = \begin{bmatrix} \mathbf{W}_{\mathrm{T,V}} & \mathbf{W}_{\mathrm{T,H}} \\ \mathbf{W}_{\mathrm{T,V}} & \mathbf{W}_{\mathrm{T,H}} \end{bmatrix}.$$

Here, $\mathbf{V}_{\mathrm{R,V}}$ and $\mathbf{V}_{\mathrm{R,H}}$ are the baseband combiner matrices for vertical and horizontal polarizations, respectively. Receive antenna weight matrices $\mathbf{W}_{\mathrm{R,V}}$ and $\mathbf{W}_{\mathrm{R,H}}$, baseband precoder matrices $\mathbf{V}_{\mathrm{T,V}}$ and $\mathbf{V}_{\mathrm{T,H}}$, and transmit antenna weight matrices $\mathbf{W}_{\mathrm{T,V}}$ and $\mathbf{W}_{\mathrm{T,H}}$ are all defined following the same logic.

At this point, the only term missing is the dual-polarized channel matrix $\mathbf{H}$. It can be expanded as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\mathrm{VV}} & \mathbf{H}_{\mathrm{VH}} \\ \mathbf{H}_{\mathrm{HV}} & \mathbf{H}_{\mathrm{HH}} \end{bmatrix},$$

where $\mathbf{H}_{\mathrm{VV}}$ and $\mathbf{H}_{\mathrm{HH}}$ are the co-polarized channel matrices for vertical and horizontal polarizations, respectively. $\mathbf{H}_{\mathrm{VH}}$ and $\mathbf{H}_{\mathrm{HV}}$ are the channel matrices for the cross-polarized components from horizontal to vertical and from vertical to horizontal polarizations, respectively. The channel matrices define the mapping between transmit and receive antennas and all transformations in between. In the dual-polarized model, the polarization purity characteristics of the antennas is also captured by the channel model [114].

Using the aforementioned expansions, Equation (15) may be written to highlight the differently polarized signals as

$$\begin{aligned} \underbrace{\begin{bmatrix} \mathbf{x}_{\mathrm{V}} \\ \mathbf{x}_{\mathrm{H}} \end{bmatrix}}_{\text{received signal}} &= \underbrace{\begin{bmatrix} {\mathbf{V}_{\mathrm{R,V}}}^{\mathrm{H}} {\mathbf{W}_{\mathrm{R,V}}}^{\mathrm{H}} \mathbf{H}_{\mathrm{VV}} \mathbf{W}_{\mathrm{T,V}} \mathbf{V}_{\mathrm{T,V}} \mathbf{s}_{\mathrm{V}} \\ {\mathbf{V}_{\mathrm{R,H}}}^{\mathrm{H}} {\mathbf{W}_{\mathrm{R,H}}}^{\mathrm{H}} \mathbf{H}_{\mathrm{HH}} \mathbf{W}_{\mathrm{T,H}} \mathbf{V}_{\mathrm{T,H}} \mathbf{s}_{\mathrm{H}} \end{bmatrix}}_{\text{wanted co-polarized signal}} \\ &\quad + \underbrace{\begin{bmatrix} {\mathbf{V}_{\mathrm{R,V}}}^{\mathrm{H}} {\mathbf{W}_{\mathrm{R,V}}}^{\mathrm{H}} \mathbf{H}_{\mathrm{VH}} \mathbf{W}_{\mathrm{T,H}} \mathbf{V}_{\mathrm{T,H}} \mathbf{s}_{\mathrm{H}} \\ {\mathbf{V}_{\mathrm{R,H}}}^{\mathrm{H}} {\mathbf{W}_{\mathrm{R,H}}}^{\mathrm{H}} \mathbf{H}_{\mathrm{HV}} \mathbf{W}_{\mathrm{T,V}} \mathbf{V}_{\mathrm{T,V}} \mathbf{s}_{\mathrm{V}} \end{bmatrix}}_{\text{unwanted cross-polarized signal}} \\ &\quad + \underbrace{\begin{bmatrix} {\mathbf{V}_{\mathrm{R,V}}}^{\mathrm{H}} {\mathbf{W}_{\mathrm{R,V}}}^{\mathrm{H}} \mathbf{n}_{\mathrm{V}} \\ {\mathbf{V}_{\mathrm{R,H}}}^{\mathrm{H}} {\mathbf{W}_{\mathrm{R,H}}}^{\mathrm{H}} \mathbf{n}_{\mathrm{H}} \end{bmatrix}}_{\text{noise and interference}}. \end{aligned} \tag{16}$$

The received signals in both polarizations may be found using an expression similar to that shown in Equation (15). This time, however, there are three terms in the sum, and the terms are obtained using regular matrix product. As highlighted in Equation (16), there are terms for the wanted co-polarized signal, for the unwanted cross-polarized signal and for noise and interference.

The dimensions of the symbols in Equation (16) are given in Table 3. The dimensions are given assuming identical configuration in both polarizations with L independent data streams. There are also M_{R} and M_{T} receive and transmit chains and N_{R} and N_{T} antenna elements at the two ends of the link, respectively. In total, there are thus $2L$ streams fed to $2M_{\mathrm{R}}$ and $2M_{\mathrm{T}}$ chains and $2N_{\mathrm{R}}$ and $2N_{\mathrm{T}}$ antennas.

Equation (16) provides a mathematical formulation of the entire beam alignment process. While it corresponds to single-user scenario, it can be expanded to account for multiple transmitters. This is done by replicating the first two product terms in the sum for each transmitter and removing their effect from the noise term. The

transmitter weight matrices $\mathbf{W}_{\mathrm{T,V}}$, $\mathbf{W}_{\mathrm{T,H}}$, $\mathbf{V}_{\mathrm{T,V}}$ and $\mathbf{V}_{\mathrm{T,H}}$, and the channel matrices $\mathbf{H}_{\mathrm{VV}}$, $\mathbf{H}_{\mathrm{VH}}$, $\mathbf{H}_{\mathrm{HV}}$, $\mathbf{H}_{\mathrm{HH}}$ are unique to each transmitter and link, respectively.

The question is how to set the weights jointly at the two ends of the link to maximize the wanted co-polarized signals while simultaneously minimizing the two unwanted terms [57], [59]. The end-to-end signals are controlled by tuning the weight matrices at both ends of the link, while the noise and interference signals may be suppressed by appropriate selection of receiver weights.

Referring back to Equation (15), knowing the received signal $\mathbf{X}$ and its own weighting matrices $\mathbf{V}_{\mathrm{R}}$ and $\mathbf{W}_{\mathrm{R}}$, the receiver tries to recover the original transmitted signal $\mathbf{S}$ [57], [59]. This is rather straightforward if the receiver knows the channel matrix $\mathbf{H}$ and both the transmit weighting matrices $\mathbf{V}_{\mathrm{T}}$ and $\mathbf{W}_{\mathrm{T}}$. $\mathbf{X}$ still includes noise and interference terms which can be minimized if $\mathbf{N}$ is known. Now, the $\mathbf{S}$ can simply be found using matrix inversion.

On the transmitting side, knowledge of the channel $\mathbf{H}$ leads to increased performance since the system may exploit all available resources [57], [59], [116]. By controlling the transmit weighting, the transmitter may also minimize co-channel interference (CCI) in multiuser scenarios [115], [117], [118]. Knowing the channels of all links, the system may balance between performance of individual links and the interference experienced by other devices. Reciprocally, knowing all the channels and transmit weights, the receiver may minimize the effect of CCI.

The aforementioned settings are a brash idealization and far from practice [57], [59]. The endpoints may share their weight matrices but the behavior of the channel and noise sources and interferers are initially completely unknown. Only some limited information may be obtained through laborious training processes. Moreover, this information is fleeting since the channel changes over time.

Table 3: Dimensions of the terms in Equation (16) assuming L parallel data streams using M_{R} and M_{T} receive and transmit chains connected to N_{R} and N_{T} antenna elements, respectively.

Parameter	**Symbol**	**Dimensions**
Received signal	$\mathbf{x}_{\mathrm{V}}$, $\mathbf{x}_{\mathrm{H}}$	$L \times 1$
Baseband combiner	$\mathbf{V}_{\mathrm{R,V}}$, $\mathbf{V}_{\mathrm{R,H}}$	$M_{\mathrm{R}} \times L$
Receive antenna weights	$\mathbf{W}_{\mathrm{R,V}}$, $\mathbf{W}_{\mathrm{R,H}}$	$N_{\mathrm{R}} \times M_{\mathrm{R}}$
Channel matrix	$\mathbf{H}_{\mathrm{VV}}$, $\mathbf{H}_{\mathrm{VH}}$, $\mathbf{H}_{\mathrm{HV}}$, $\mathbf{H}_{\mathrm{HH}}$	$N_{\mathrm{R}} \times N_{\mathrm{T}}$
Transmit antenna weights	$\mathbf{W}_{\mathrm{T,V}}$, $\mathbf{W}_{\mathrm{T,H}}$	$N_{\mathrm{T}} \times M_{\mathrm{T}}$
Baseband precoder	$\mathbf{V}_{\mathrm{T,V}}$, $\mathbf{V}_{\mathrm{T,H}}$	$M_{\mathrm{T}} \times L$
Transmitted signal	$\mathbf{s}_{\mathrm{V}}$, $\mathbf{s}_{\mathrm{H}}$	$L \times 1$
Unwanted signals	$\mathbf{n}_{\mathrm{V}}$, $\mathbf{n}_{\mathrm{H}}$	$L \times 1$

Nevertheless, all beam steering applications involve channel estimation [57], [59] at least in an abstract sense. This is evident in systems that involve digital beamforming but is also present in single-receiver systems. The challenge is how to recover the original signal with limited information of the transformations it has undergone. Restricted by the channel and system constraints, one must find a suitable compromise between performance and complexity – especially in the multiuser scenario.

The channel forms the entire basis of beam steering, and characterizing its behavior is crucial to the discussion [45], [119]–[122]. Channel models capture both the limitations and possibilities of the radio environment while also providing a test bench for algorithm development. This section focuses on characterizing the propagation phenomena; beam steering functionality and existing approaches are further discussed in Sections 5 through 7.

Employed in mobile communications for more than thirty years[3], channel behavior at lower frequencies is well-known [48], [86], [88], [89], [123]. At mmW frequencies, however, the work related to channel characterization especially in outdoor environments is still underway with new findings being published regularly [45], [46], [119]. Next, the key mmW propagation phenomena are discussed.

3.4 Multipath effects

The radio channel behaves quite differently at millimeter waves compared to conventional frequencies as highlighted in Figures 19 and 20 [72]. The figures illustrate the simplified impulse response of the radio channel in the delay-angle domain, normalized to the LOS component. The figures are drawn for the same exact deployment scenario assuming a LOS path and a number of non-line-of-sight (NLOS) paths due to objects in the surrounding environment.

Figure 19 clearly shows multipath fading at lower frequencies. The situation corresponds to Rice distributed channel, meaning there are a LOS component and a large number of multipath components distributed more or less evenly in the entire delay-angle plane [84], [86], [87]. This is not the case at mmW frequencies, as is evident from Figure 20. While there are still both LOS and NLOS components present, the delay-angular spread is completely different. This time the LOS component is clearly dominant, and the multipath components are grouped in both delay and angle domains [72], [121], [124]–[127]. Referred to as *clusters* and marked with shaded circles of different colors, each group consists of a small number of impulses of similar power, delay and angle of arrival.

Only the azimuth angle is shown in Figures 19 and 20. At lower frequencies, the situation can usually be considered as two-dimensional; there is no real need to consider the elevation angle as required gains are achieved with omnidirectional or sectorized antennas [48], [84], [86], [87]. At mmW frequencies, however, three-dimensional world is typically considered as the antenna is directive in both azimuth

[3]First-generation *Nordisk Mobiltelefoni* (NMT) systems were first launched in 1981 [53].

and elevation planes [47], [48], [119], [128]. The situation in the elevation plane is similar to that in the azimuth plane but the angular spread is smaller and the NLOS components tend to concentrate on elevation angles above the LOS component [47].

Figures 19 and 20 highlight the major changes when moving from conventional frequencies to millimeter waves. While a crude overview was already given above, an in-depth discussion about different phenomena is given next. As is typical at mmW frequencies, the discussion considers only links with high-gain antennas at both ends of the link [32], [38], [44], [46], [49]. As a result, some of the phenomena may be attributed not only to smaller wavelength but also to greater directivity of the antennas employed.

Millimeter-wave communications are LOS dominant [46], [121], [124], [129]. NLOS components experience 10 ... 50 dB attenuation compared to the direct LOS path. The best NLOS paths are typically 20 dB weaker than the primary LOS link. This setting is challenging for link budgeting if NLOS paths are considered. Due to this severe attenuation, the NLOS range is typically limited. The NLOS range cannot be increased solely by means of transmit power control for two reasons. Firstly, the required range of power control is very wide, and secondly, there are rules and regulations regarding maximum transmit power as already discussed in Section 3.1.

LOS dominance can be attributed to both the two aforementioned causes. First, multipath components are not direct but propagate using objects in the surrounding environment. These objects are not properly illuminated by the antenna since directive antennas focus the radiated energy in the main lobe [54], [55], [62]. As a result, the indirect components are weaker to begin with as antenna gains drop quickly even at small angular offsets. The NLOS coverage may, however, be improved within a smaller area by using a smaller gain antenna [130].

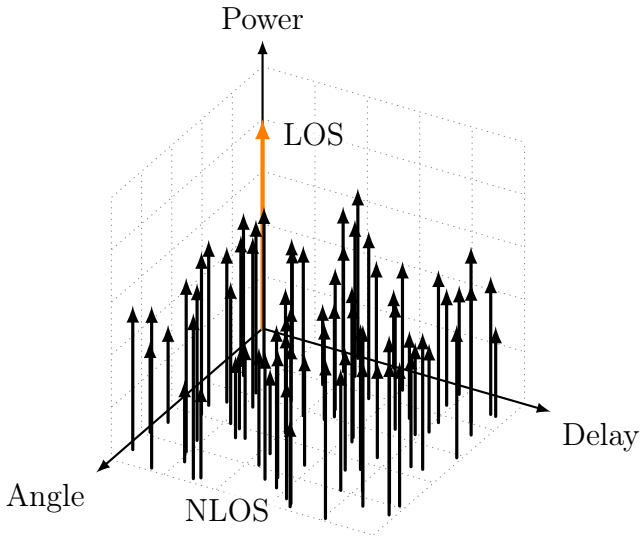


Figure 19: Delay-angle distribution at low frequencies.

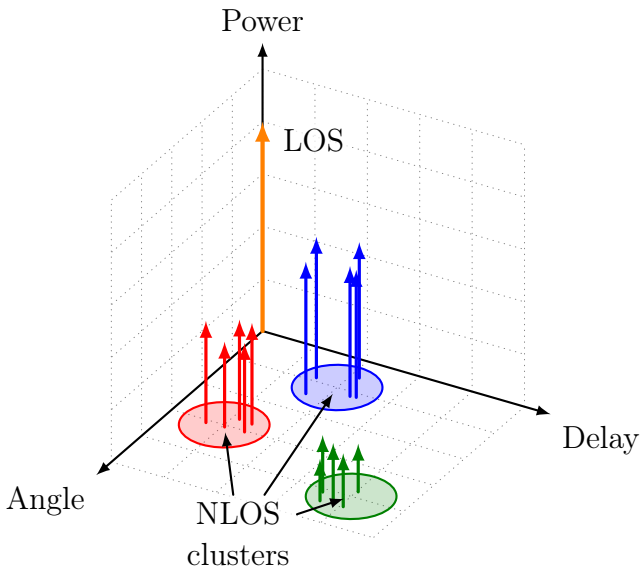


Figure 20: Delay-angle distribution at millimeter waves.

Second, the paths experience higher loss not only due to greater distance but also as they interact with the objects [46], [84], [86]. These interaction losses tend to increase when the wavelength decreases as the objects get larger compared to the wavelength [86]. Moving from conventional frequencies to millimeter waves, all dimensions are effectively scaled up in same relation as the wavelength decreases. The situation may thus be considered quasi-optical [131].

Given the small wavelength of only 3 . . . 5 mm, all objects seem gigantic, and even surface roughness of buildings and pavement is comparable in size. Consequently, penetration losses are increased due to greater penetration depth measured in wavelengths [46], [86], [132]. Furthermore, diffraction losses are also increased as edges are no longer sharp [86], [119], [132]. Finally, increased surface roughness results in scattering where radio waves reflect in arbitrary directions. As a result, Wyne *et al.* [132] and Räisänen *et al.* [133] suggest considering only first-order reflections as viable NLOS paths.

Compared to the situation at conventional frequencies, not only are the multipath mmW components significantly weaker in power but also smaller in number [67], [76], [113], [127], [134], [135]. The environment is considered *sparse* and the components are clustered in both delay and angular domains instead of being approximately evenly distributed. Moreover, the signals tend to be less spread in either domain when either distance or antenna directivity is increased [46], [47], [121], [136]. This clustering behavior has a number of consequences which are important to this application.

First, the behavior depicted in Figure 20 clearly shows the need for beam alignment [37], [47], [121], [137]. The situation can be described as if the channel had severe fading that is made even worse if antenna gain is increased [125]. There are only a handful of different directions where the signal quality is good enough for a reliable

link [47], [67], [127]. Even in these directions high gain is required due to larger path loss whereas the signal strength in other directions remains insufficient for reliable operation.

Second, high attenuation is beneficial with co-channel interference (CCI) in mind [51], [121], [127], [138]. Different devices operating at the same frequency are not easily disturbed as the interfering signal levels are limited. Consequently, the capacity of millimeter-wave links tends to be limited by noise rather than by interfering signals. As already discussed in previous subsection, interference toward others may be minimized by controlling transmit power [115], and hence the interference-limited case is also referred to as being power-limited [121].

Third, there are only a small number of viable paths to connect the receiver and transmitter [67], [125], [127]. The mmW channel is thus said to be of *low rank* [52], [67], [127]. Luckily, it still has enough spatial degrees of freedom to support a small number of parallel data streams, and low-order spatial multiplexing, i.e., MIMO processing is possible [67], [110], [127]. For instance, Akdeniz *et al.* [127] suggest using up to three data streams simultaneously. However, performance improvements are achieved only if the SNR is high enough to be able to split a single beam into multiple beams [67], [127]. It is similarly possible to use multiple beams as a diversity technique to improve availability [124], [127], [136].

Finally, while channel is time-variant, the paths are reciprocal [67], [69], [113], [139]. Moreover, the mmW channel maintains the polarization of the transmitted signal over LOS links [129]. Some polarization mixing is evident in NLOS schemes due to interaction with the environment [129], [131]. For instance, the handedness of circularly polarized signals is changed upon reflection [129], [131]. Furthermore, degradation due to polarization mismatch of up to 20 dB can be expected [75], [131], suggesting for polarization diversity [67], [131].

This polarization purity is commonly measured using cross-polarization discrimination (XPD) defined as the ratio of the co-polarized and cross-polarized powers, P_{C} and P_{X}, respectively [140]:

$$XPD = \frac{P_{\mathrm{C}}}{P_{\mathrm{X}}}. \tag{17}$$

In LOS dominant mmW scenarios, polarization mixing is negligible and the XPD is typically 25 ... 30 dB [129]. However, Thomas *et al.* [47] suggest using an XPD of 15 dB in mmW channel characterization to account for multipath effects and imperfectly aligned beams.

Effectively limiting the SNR, an XPD of 15 dB can reduce the performance of some co-channel dual polarization (CCDP) systems [87], [99]. Cross-polarization interference (XPI) is somewhat similar to ISI since it cannot be reduced with adequate power control. Instead, one can choose the used polarization with XPD characteristics in mind or resort to a more robust modulation with smaller spectral efficiency and throughput. For instance, circular polarization is immune to rotation and is favored in satellite communications affected by Faraday rotation in the ionosphere [86], [123].

However, circular polarization and especially dual circular polarization complicates the antenna design significantly [54]–[56], [62].

Additionally, the effect of XPI can be reduced with appropriate selection of used polarization cross-polarization interference cancellation (XPIC) techniques [87], [99], [101]. While these methods achieve significant improvements, they can induce non-negligible branching losses while also increasing hardware complexity and bill of materials considerably.

Regardless, CCDP systems have been suggested as a viable means for increasing throughput [47], [67], [110], [114], [115] of mmW radio links. This is reasonable given the need for high beamforming gain and the low rank of the channel. Moreover, both polarizations are typically aligned at the same time, reducing the training time [114], [115]. However, the differently polarized radio waves may still exhibit distinct behavior. For instance, the reflection coefficients given by Fresnel equations are dependent on the polarization [141].

3.5 Channel modeling

Having introduced the key propagation characteristics at mmW frequencies, we continue to consider actual channel models that capture these phenomena. At lower frequencies, a multitude of channel models exists with an emphasis on different scenarios or goals [48]. None of these models meets the requirements for next-generation mobile communications systems [48]. For example, the models do not support frequencies up to mmW, gigahertz bandwidths or antenna arrays with hundreds of elements. New models are thus required.

Examples of notable low-frequency channel models include ETSI/3GPP (European Telecommunications Standardization Institute and 3rd Generation Partnership Project, respectively) SCM (spatial channel model) for MIMO simulations [142], 3GPP 3D-UMi and 3D-UMa (three-dimensional urban micro and urban macro) models [143], WINNER [144] and WINNER II [145] models by the IST-WINNER (Information Society Technologies – Wireless World Initiative New Radio) group, WINNER+ [146] by the CELTIC (Cooperation for a sustained European Leadership in Telecommunications) project, COST2100 [147], and those defined for third-generation IMT-A (International Mobile Telecommunications-Advanced) systems [148] by ITU-R, commonly referred to as ITU-R UMi and UMa models.

At millimeter waves, there are no mainstream models as the work is still underway [47], [119], [120], [122], [128]. There only a small number of proposals that have started to gain attention. Furthermore, most of these models, such as the mmW extension of the Saleh-Valenzuela model [149] and the IEEE 802.11ad channel model [150], are intended for short-range indoor communications. While they essentially capture the same phenomena, the transition from indoors is not straightforward [52], [120], [125]. This is because all the dimensions, materials, networking equipment and goals change dramatically, and weather effects come into play. Moreover, the proposed

models typically focus on the co-polarized case for simplicity reasons. Regardless, such models may still be used in some situations with modest requirements.

Broadly speaking, there are two approaches to channel modeling: *deterministic* and *stochastic models* [46], [48], [151]. While deterministic models are based on ray tracing principles or full-wave electromagnetic (EM) simulations based on Maxwell's equations, distribution functions form the heart of stochastic models. Stochastic models are typically favored for their simplicity and ease of use. In fact, all the models listed above have been stochastic models. On the other hand, deterministic models are not only more straightforward to develop but also typically more accurate. In the following, examples of both types of models are presented.

First, we discuss deterministic ray tracing models, sometimes referred as *physical channel models* [151]. The channel response is obtained by launching a ray from the transmitter in all directions, covering a full three-dimensional sphere or two-dimensional circle with some angular resolution [46], [151]–[155]. Each ray travels in a straight line but it can also reflect, diffract or scatter from, or penetrate through various objects in the environment. These rays are followed or *traced* to determine if they are able to reach the receiver. It is also possible to start from the RX and trace the rays back to the TX.

Although radio waves are spherical electromagnetic waves rather than rays, rays capture their behavior effectively [46], [48], [151]–[154]. Ray tracing has all the makings of an accurate model but only if the deployment environment is properly modeled. After all, the behavior of the rays is dependent on the surrounding environment – more specifically on its geometry and electromagnetic properties of the used materials. Hence they are also referred to as *site-specific models* [152], [156].

Accurate ray tracing thus actually concerns accurately modeling the environment [46], [151], [152], [154], [156]. While replicating the environment is conceptually easy and straightforward in practice, the approach has numerous drawbacks. First of all, building the model is extremely tedious and time consuming, resulting in simplifications that lower accuracy. Consequently, the resulting model represents only the single scenario with given objects and their placement, including the locations of the networking equipment.

Having built an accurate three-dimensional model of the environment once, the same skeleton can be used with different TX/RX placement by running the simulation multiple times [46], [48], [151], [154]. While such arrangement is possible, it is also very time consuming and it does not increase the model coverage beyond the specified instances. Aiming for just one degree resolution in both azimuth and elevation planes, there are over 64,000 rays to trace per simulation.

Because of the large number of rays, the models typically include strict break conditions and other simplifications [46], [48], [152], [154]. For instance, only a very limited number of object interactions are associated with each ray. For instance, only two reflections may be permitted, and only one global value for reflection loss is used. Furthermore, some propagation mechanisms such as scattering, penetration and/or

diffraction may be completely excluded from the model. Moreover, any rays that escape the modeled area are considered lost, and the simulation continues by tracing the next ray.

While deterministic models are essentially about replicating the geometry of environment, stochastic models take quite a different approach instead [46], [151]. Also known as statistical models, they are based on modeling the channel response mathematically through statistics, i.e., distribution functions. In other words, the model includes a response function with an arbitrary number of parameters with values determined using predefined distribution functions.

Tapped delay line is a typical approach to stochastic channel modeling [88], [118]. For modeling millimeter waves in a street canyon environment, Figure 20 provides a good starting point for the response function. The clustered multipath components can, for instance, be thought to have been generated by a number of distinct scatterers [46], [47], [120]. The cluster-wide parameters, i.e., delays and both azimuth and elevation angles of each group, may for instance be obtained based on their location that is randomized with certain criteria [46]. Within each cluster, properties of individual rays are then found using a distribution function.

Not limited to a specific location, stochastic models are more general than the deterministic counterparts [46], [151]. Other benefits of stochastic modeling include their ease of use and low computational complexity. While they are applicable to a number of different scenarios, stochastic modeling is less accurate than ray tracing. Furthermore, stochastic models are challenging to come up with as determining the parameter distributions involves extensive measurements and data analysis. Hence, they are occasionally referred to as empirical models.

The required measurements are far more complex or demanding than simply determining the geometry of the modeled location [46], [151]. For accurate ray tracing, material parameters are also needed. However, these are often already available and cataloged [53]. Such a catalog and a few good photographs of the location may very well suffice. The distribution functions, on the other hand, are typically determined by carrying out channel sounding measurements in a realistic deployment environment.

The sounding measurements are typically done as function of various parameters such as frequency, polarization, placement of the networking equipment and type of antenna [46], [47], [130], [157]. These full-fledged radio frequency measurements require not only expensive instruments but also skilled labor, precise calibration and proper procedure if meaningful results are to be obtained. Moreover, the channel parameters are typically correlated which needs to be accounted for [47], [48], [88], [148].

Even though stochastic models are not location specific by themselves, there are some restrictions due to the experimental nature behind the distribution functions [46], [155]. The measured distributions are typically valid only in a similar scenario, and the model can predict the behavior accurately only in the characterized setting. As

a result, stochastic model suites – like those listed above – typically include multiple sets of parameters corresponding to different deployment environments.

It is also possible to combine the aforementioned approaches in a number of ways. For instance, the entire model can be derived using ray tracing simulations instead of carrying out extensive measurement campaigns [156], [158]. Another approach is to characterize the channel deterministically while the scattering properties and cluster locations are obtained stochastically [46], [128]. Referred to as geometry-based stochastic channel modeling (GSCM), the aforementioned COST2100 [147] and WINNER [144] models are for instance based on this concept.

Next, two of the more recent channel models that are applicable to the situation of interest are introduced. We start with the model promoted by the METIS project [48]. Within the project, three channel models are defined: a deterministic model, a stochastic model and a hybrid between the two. The models address the requirements, specifications and goals [8], [9] defined within the project for upcoming mobile communication systems. Together, the models form one of the most notable channel modeling frameworks to date.

Only the deterministic model, or the *map-based model* as it is referred to within the project, is discussed here. This is because the map-based model is most comprehensive of the three, and most importantly, because it includes full mmW support unlike the other two. Developed from the aforementioned WINNER and 3GPP models, the stochastic METIS model mainly concerns lower frequencies below 6 GHz. The model includes only a few propagation scenarios around 60 GHz, most of which are indoor settings.

The map-based model is perhaps the most comprehensive model to date. Defined from 450 MHz to 100 GHz, it accounts for all important phenomena in the radio channel using ray tracing. For instance, the mode includes support for a multitude of different propagation mechanisms and beamforming with MIMO, providing accurate results in real-world scenarios with moving user in a dynamic environment. It covers all propagation scenarios defined within the METIS project [8], [9] except those defined for rural macro setting. The reason for this is most likely the challenges associated with modeling a generic rural environment with wide open spaces and lots of vegetation.

The model consists of four main phases. At first, the environment is created including various objects, scatterers and TX/RX placement. For street canyon environment, or *dense urban information society* as it is referred to within METIS, the Madrid grid model introduced in Section 3.2 can be used. In the second phase, interaction points, propagation pathways and their characteristics, such as path length and directions at both ends of the link, are determined.

Interaction points (IPs) refer to locations in which the radio waves interact with the environment. They are also referred to as secondary nodes since new multipath components are generated as a result of the interaction. Within the channel model, there are two kinds of IPS: diffraction edges and scattering sources. Reflection is not

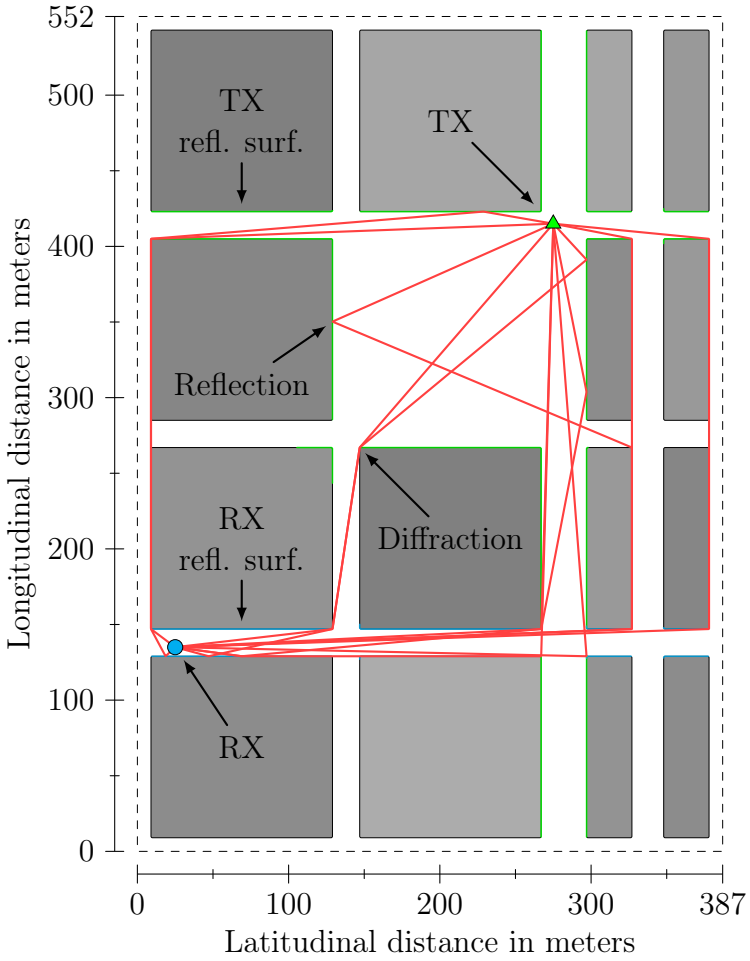


Figure 21: Examples of possible propagation paths considered by the METIS map-based model [48].

modeled using IP but rather the points are found using specular reflection. That is, only the IPs that are visible either directly or via a single reflection from either the TX or RX are considered. Consequently, the number of reflections within any path is limited to two.

In the next phase, propagation pathways then are divided into segments with different propagation phenomena as is highlighted in Figure 16: direct propagation or interaction with objects in the environment. For simplicity, only paths with 5 or less segments are recommended. The channel matrices of these segments are determined based on the propagation type in said segment. For instance, diffraction is defined for both vertical and horizontal edges using either the Berg's model or uniform theory of diffraction (UTD). The effect of scattering is found using surface roughness, while a simple screening procedure is used for shadowing. Some examples of possible propagation pathways and the reflection surfaces are shown in Figure 21.

Finally, having determined the polarization mixing, shadowing loss, arrival and departure angles and path lengths, the transfer function between the endpoints is

determined. During the last phase, the channel impulse response is determined, also accounting for antenna characteristics. It should be noted, however, that large antenna arrays should be characterized as individual antennas during the first phase.

Even though not as complex as full-fledged ray tracers, the main drawback of the map-based model is its complexity. There are, however, a number of different ways to find a suitable compromise between accuracy and complexity. For instance, it is possible to consider only a static environment, use simplified diffraction modeling, reduce the dynamic range and remove some of the objects with low impact on the result. Even at the most accurate setting, the model is unable to capture some phenomena present in measurements. This is because of the modeling accuracy; it would require immense effort to model all small features in the environment.

Our discussion on the model has been necessarily short, abstract and more of an overview. Complete descriptions of the three METIS channel models and step-by-step instructions are available in *METIS Deliverable D1.4 – METIS Channel Models* [48]. It is important to note that no public implementations of the model exist at the time of this writing.

Finally, we consider a stochastic mmW channel model proposed by Thomas *et al.* [47]. This GSCM is also the model of choice within the MiWaveS project, probably to due to its simplicity. The model is intended for E-band transmission in an urban micro setting in both LOS and NLOS conditions with hop lengths less than 200 m. Similar to the aforementioned 3GPP 3D-UMi model [143], the proposed model follows the same methodology and format but with much higher carrier frequency and bandwidth.

The model was created by combining and fitting together ray tracing simulations and real-life measurements. Comprising a number of different small cell access link scenarios, the wideband channel sounding was carried out at 73 GHz in New York City. Due to limitations in the measurement setup, the measurement environment was also modeled in a ray tracer and simulation results were used to extend and complement the empirical results. While reflection is considered as the primary propagation mechanism, the simulator also accounts for diffraction but not scattering. The interaction losses were found using simplified empirical framework instead of using complex but established theory.

The channel model includes direct LOS component and five clusters with ten rays each. The parameters of individual rays are generated either using distribution functions or scalar coefficients provided in the proposal. For instance, the azimuth and elevation angle distributions are wrapped Gaussian and Laplacian functions, respectively. Determining channel transfer function involves the following five stages. First, the omnidirectional path loss PL is found similarly to ITU-R UMi model [159] as

$$PL_{\mathrm{dB}}(r) = 20\log_{10}\left(\frac{4\pi\cdot 1\,\mathrm{m}}{\lambda}\right)\,\mathrm{dB} + 10n_{\mathrm{PL}}\log_{10}\left(\frac{r}{1\,\mathrm{m}}\right)\,\mathrm{dB} + \sigma_{\mathrm{SF,\,dB}}, \tag{18}$$

where λ is the wavelength, n_{PL} the path loss exponent, r the link length and σ_{SF} the shadowing factor. Thomas *et al.* suggest using $n_{\mathrm{PL}} = 2.1$, $\sigma_{\mathrm{SF}} = 4.9$ dB and $n_{\mathrm{PL}} = 3.3$, $\sigma_{\mathrm{SF}} = 7.6$ dB for LOS and NLOS links, respectively.

In the following stages, the cluster-wide properties are first generated before determining properties of individual rays. In the second phase, cluster delays are determined; the cluster delays are exponential while inter-cluster delays are grouped. Next, cluster powers are determined and normalized. Finally the angular spreads are generated using aforementioned distribution functions. A more detailed description of the channel model is given in the original proposal [47]. This also includes step-by-step instructions and tables for the various parameters.

4 Target system

This section considers the envisioned next-generation network architecture, including both the node functionality and the system as a whole. The goal is to provide some background information while also highlighting various aspects to consider when discussing the beam alignment functionality. From the system architecture and organization, we can devise both requirements and constraints for the alignment application. However, the actual alignment functionality is dependent on the beam steering hardware available in the individual nodes.

The target system described here corresponds to the one promoted by the MiWaveS project [34]. The system-level aspects of the system are discussed in Section 4.1. This subsection focuses on the interplay between different networking devices rather than the operation of individual nodes. The operation of individual nodes is discussed in Section 4.2 using the MiWaveS demonstration hardware [160] as an example implementation. The hardware is based on an instrumentation platform developed by National Instruments [161] discussed in Section 4.3.

4.1 Network organization

Already outlined in Section 1 and illustrated in Figure 1, the next-generation mobile networks will involve numerous mmW small cells to supplement low-frequency macro coverage [7], [12], [28], [33]. Having no wired connection, these small cell APs will work as autonomous wireless relay stations. The mmW APs will organize themselves into a mesh network, overseen by the base station. These meshes connect to the core network via a gateway (GW) at the BS or at an AP where additional wired connectivity is available.

Employing multiple air interfaces and comprising multiple self-organized mesh networks, the network is heterogeneous in structure. These meshes are separate and the network comprises multiple smaller subnetworks rather than a single, all-covering mesh. While a complete mesh would be more robust toward gateway failures, the advantages of subnets are overwhelming in comparison [162]–[164]. Even the Internet is hierarchical, organized into subnetworks.

Considering our system, some of the major advantages of subnetworking may be outlined as follows [162]–[164]. The greatest advantage is the ease of network management; a separate configuration can be run in each subnet. Second, fewer devices results in fewer compatibility issues. Third, contagious errors are by default limited to nodes in the subnet. Fourth, clock skew across a smaller network is not as significant [165]. Fifth, finding a reasonable schedule is easier if there are fewer competing tasks [166]. In terms of the target system, these "tasks" correspond to active links.

Here we focus on *time-division duplexed* (TDD) systems where both TX and RX use the same frequency band but not at the same time [83], [84], [88], [89], [99], [167], [168]. TDD is favored over frequency-division duplexing (FDD) for its flexibility

due to four main benefits. First, TX-RX separation may be accomplished using a simple switch without costly diplexers. Second, hardware is further simplified since dual-band operation is not required. Third, decided by means of time allocation, uplink-downlink ratio is dynamic. Fourth, network organization requirements are relaxed since frequency planning is not needed.

Another significant benefit of TDD systems is that they are always *reciprocal* [88]–[90], [99], [100], [168]. A wireless system is reciprocal if the transmission characteristics are identical for both TX and RX. Reciprocity is a beneficial feature since the same set of weights can be used for both transmit and receive. This applies to both beamforming and baseband MIMO combining.

On the other hand, the half-duplex nature has its shortcomings [84], [99]. First, the two ends of the link have to take turns transmitting and receiving. Consequently, the endpoints need to be properly synchronized and scheduled to avoid situations where both nodes are transmitting or receiving at the same time. Secondly, TX-RX coupling may cause problems in multibeam, i.e., multitransceiver, nodes. That is, there might be restrictions concerning simultaneous transmit and receive using different beams. Third, associated delays are larger than in corresponding FDD systems. Fourth, efficient feedback organization is difficult in TDD systems, and limits the performance of some alignment algorithms.

The challenge in feedback organization can be attributed to the substantial overhead associated with TX-RX role reversal. This concept is illustrated in Figure 22 where a timing diagram of a TDD link is shown. The diagram depicts a situation where the primary operation, i.e., beam training, is interrupted to feed the results back to the TX. In this case, the roles are reversed twice, and each reversal causes a *fixed* delay independent of the length of the turn. During this delay, the endpoint initially transmitting remains effectively inactive and simply waits for the receiving end to start transmitting.

This role reversal delay consist of a *propagation delay* (less than 330 ns for distances up to 100 m) [86], [123], *processing* or a *decoding latency* [169], [170] (typically dominant of the three), and a *transition guard* [88], [99]. Used to account for signal settling and any offset in synchronization, the guard time functions similar to a guard band but in time domain. A rigorous analysis would also include an encoding delay and a beam alignment delay but for simplicity these are assumed

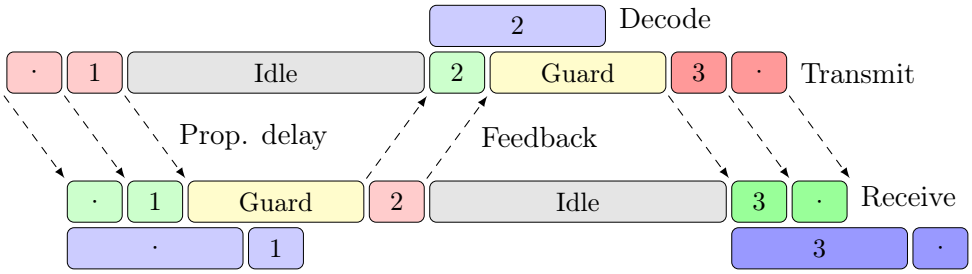


Figure 22: A timing diagram of a feedback cycle in a half-duplex system.

to be negligible compared to the decoding latency. Finally, there is a *transmission delay* due to actually having to transmit or receive the information.

Feedback is available at the receiver only after the lengthy feedback cycle is completed, starting from the slots with a shaded background. While a dialog-like setting would be best for the two-ended beam alignment process, implementing one-to-one signaling using TDD is not really a viable alternative. Instead, the role switching should be minimized. On the other hand, this increases the feedback delay which has a negative impact on the performance of preemptive or dynamic approaches. One should thus find a suitable balance between the switching delay and the overshoot due to delayed feedback. In other words, a small number of beams can be tested consecutively in a *burst*, followed by a cumulative feedback.

In case of a reciprocal full-duplex system, no designated feedback would be needed at all. If the conditions are identical for both transmit and receive, the endpoints may simultaneously measure the signal quality. Since the quality is identical at both ends, the endpoints may independently very the connection. This simultaneous measuring is elaborated in Figure 23 showing an FDD system capable of transmitting and receiving at the same time. While propagation or processing delays cannot be avoided, the link is never inactive. The training can continue without interruptions until a suitable beam pair is confirmed, and the link automatically starts using it.

While the approach described above certainly reduces latencies and improves performance, FDD systems employ different frequencies for uplink and downlink, and cannot thus always be assumed to be reciprocal [88]–[90], [99], [100], [168]. If the system is nonreciprocal, the uplink and downlink need to be aligned separately. However, if supported by the hardware, the links can be aligned simultaneously in parallel but now the alignment needs to be confirmed via feedback similarly to a half-duplex system.

Another possibility is to use a secondary low-frequency connection dedicated to exchanging feedback and other low-rate control information. Separate from the mmW link, the training process need not be interrupted to give feedback. Instead, the control channel works in parallel with the mmW link and the training may be

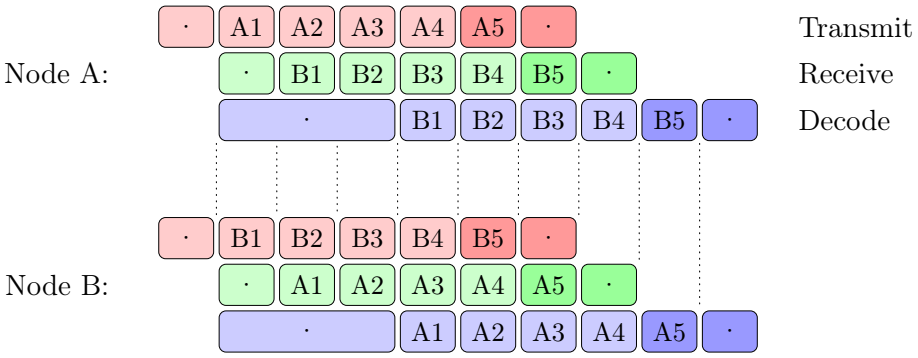


Figure 23: A timing diagram of a reciprocal full-duplex system.

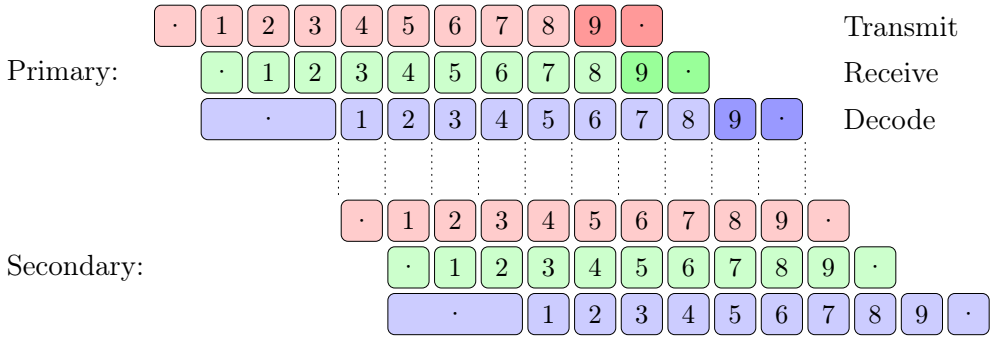


Figure 24: A timing diagram of a half-duplex system with separate feedback connection.

pipelined similarly to full-duplex systems. A timing diagram of such an arrangement is shown in Figure 24.

While separate control connection clearly reduces the feedback overhead on a half-duplex mmW link, a full cycle is still needed. While the *latency* associated with a single feedback improves only a little, the feedback *throughput* is greatly improved. To keep the latency tolerable, the processing delays associated to this control link must be minimized. The used coding scheme should thus be easy to encode and decode, but most importantly, also this link needs to be direct. Current cellular technologies cannot be used as such since *ad-hoc*, i.e., device-to-device, connectivity is not supported [83], [84], [99], [100]. Luckily, 5G system will most likely feature ad-hoc communications [10], [16].

In addition to separating uplink and downlink, also different users of the radio resource need to be multiplexed [83], [84], [89], [90], [100], [118], [163]. In the envisioned system, there are two distinct groups competing over a shared resource. In addition to forming multiple links between an AP and its users, an AP will also be connected to a number of other APs in the same mesh network. If both the backhaul and access link use the same radio – referred to as *in-band* backhaul – all links compete over the same resource. In this thesis, however, it is assumed that mobile access and backhaul are implemented using their own radios and are thus completely independent of one another.

The different links can share the resource if they are separated in time, in frequency or use unique codes to encode the message. These basic multiple access schemes are known as *time-division*, *frequency-division* and *code-division multiple access* methods (TDMA, FDMA and CDMA), respectively. In multiple-antenna systems the users may also be identified by their location, or rather their effective radio channel. This approach is commonly known as *space-division multiple access* methods or SDMA for short.

Regardless the access scheme, different users are typically found in different directions and thus typically use distinct beams. As a result, up to as many transceivers as there are *simultaneous* users are required. However, if the users were *interleaved*,

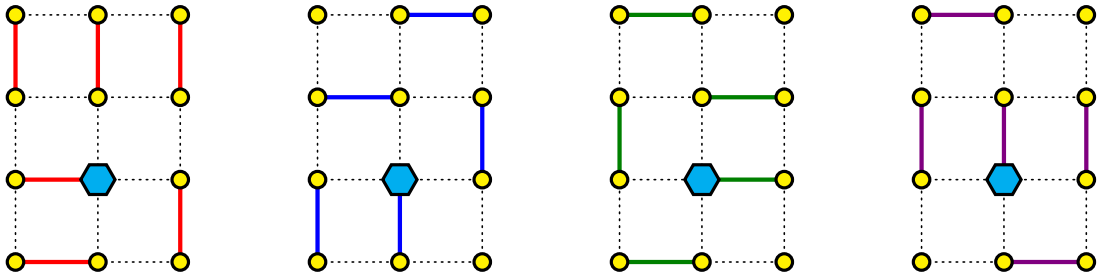


Figure 25: Four transmission sets showing active links.

a single transceiver would suffice. Consequently, separating users in time is the only practical choice. While time-division increases delays and requires synchronization, it is also very flexible and can be implemented using simple hardware [83], [84], [100]. In fact, TDMA is the only option for single-transceiver nodes. But if a node supports multiple independent beams, SDMA principles may be directly applied to communicate with multiple users simultaneously [89], [90].

While access links may be locally scheduled by the AP in question, the backhaul links are network-shared resources. As a result, the links need to be globally scheduled [166]. In the target system, this is done using a sequence of *transmission sets* repeated over time [171]. These global and semi-permanent transmission sets determine the simultaneously active links during the scheduling slot. While single-transceiver node may not connect to multiple active links at a time, crowded links may be included in multiple sets. The two-node links within a single transmission set are disjoint whereas different transmission sets need not be.

An example of a such a network schedule is given in Figure 25. In this example, one gateway node and eleven relay nodes marked with a cyan hexagon (⬡) and yellow circles (○), respectively, are scheduled using four transmission sets drawn using different colors. This schedule determines only the active link and nothing more. The scheduling slot is further divided into TDD slots, locally and dynamically assigned for uplink or downlink based on current traffic [99]. It should, however, be noted that to avoid possible TX-RX coupling, this assignment can also depend on other links if one or both of the nodes support simultaneous beams.

Focusing on the mmW functionality, the term *network* is henceforth used to refer to a single time-division mesh subnetwork. Given the strict delay constraints of 5G systems, the maximum number of hops is limited to three or so [171]. The number of other nodes and users an AP may connect to is similarly limited. This *node degree* or *fan-out* limitation arises from network scheduling since a single-transceiver node may transfer data only with a single endpoint at any one time. As a result, the number of APs is limited to no more than some tens with less than a hundred wireless links.

When discussing the beam alignment functionality in upcoming sections, following assumptions are made. First, no mmW links are formed between two floating nodes.

Instead, at least one end of the link is synchronized and has access to all relevant information on network organization, such as scheduling. Second, contrary to floating nodes, all nodes part of the existing network are properly synchronized and completely aware of the network. That is, they have an up-to-date configuration, including information regarding other nodes, network topology, schedule and routing information. Third, while METIS defines direct inter-UE connectivity, ad-hoc mmW links are ruled out. Fourth, the alignment process is disruptive, and all relay operations are put on hold for the duration. Finally, the subnet may be constructed using dissimilar devices that meet certain criteria.

4.2 Node functionality

Next, the operation of individual network nodes are discussed by studying the MiWaveS demonstration/prototyping hardware [160] as an example. We start by defining the necessary functionality on a high abstraction level. This is followed by a discussion on the key radio frequency components, and how they are used to implement a functional network node.

Figure 26 shows the high-level functional blocks found in a cellular radio transceiver [83], [84], [99], [100], [167]. In the figure, a short description of each block is also given in terms of the prototyping system. Uninterested in networking aspects, this thesis focuses on link functionality. For this reason, the network is abstracted almost completely, and the *Open Systems Interconnection* (OSI) model [172] is only considered up to the second layer. The prototype thus incorporates only physical and data link layer functionality.

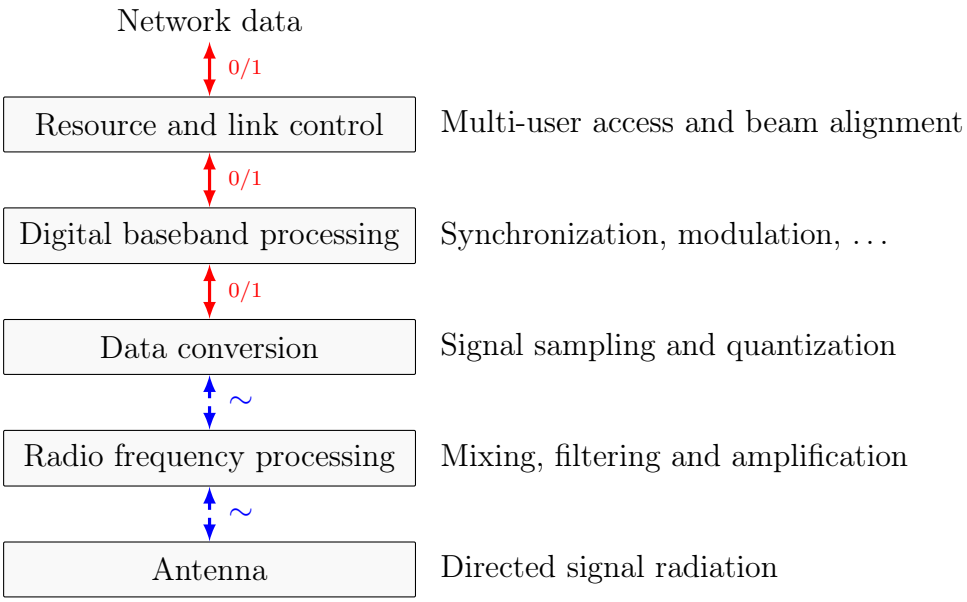


Figure 26: Functional elements in the prototyping transceiver.

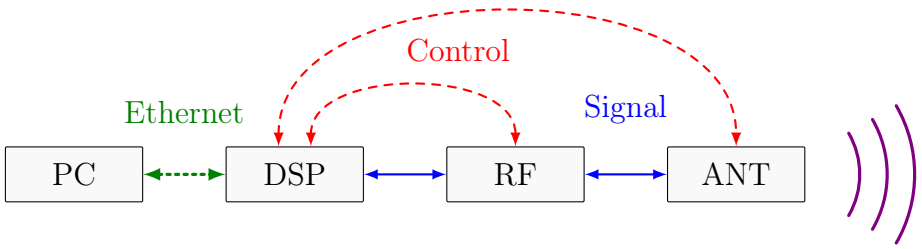


Figure 27: Prototyping setup.

Starting from the top, the first functional block is resource and link control responsible for multi-user access and beam alignment. In terms of the example system, this block corresponds to the data link layer functionality. Next, there is digital baseband processing in the physical layer. The associated tasks include time and frequency synchronization, channel estimation and equalization, modulation and coding, pulse shaping and re-sampling.

Toward the air interface, there are blocks for data conversion, analog processing at radio frequency (RF), and the antenna. Data converters transform between digital and analog domains through signal sampling, quantization and generation. RF processing involves frequency shifting between the baseband (BB) and the carrier frequencies, signal filtering and amplification. The antenna functions as the air-interface, radiating the signal in wanted directions.

The prototyping setup is built around National Instruments (NI) instrumentation system using a digital signal processing (DSP) platform controlled with a computer. Using the setup shown in Figure 27, the link achieves 2100 Mbps throughput. The node comprises an instrumentation computer (PC) connected to the DSP unit via an Ethernet connection. The main signal path connects both the transceiver module (RF) and the antenna (ANT) to the DSP unit. The DSP unit is also connected to the other modules via digital control connections. Full technical specifications are given in Table 4.

Table 4: Technical specifications of the prototyping system.

Parameter	**Value**
Operating frequency	71 ... 76 GHz
Bandwidth	1000 MHz
Throughput	2100 Mbps
Carrier division	SC-FDE
Modulation scheme	16-QAM
Coding rate	7/8
Duplexing method	Time-division
Multiple access method	Time-division

4.3 NI DSP platform

At the heart of the entire prototyping system is the National Instruments DSP unit, programmed and controlled using a tool chain by the same provider. The unit is built on top of a *NI PXIe-1085* PXI Express (PXIe) chassis [173], used to interconnect various application-specific PCI eXtensions for Instrumentation (PXI) Express compliant modules. The chassis includes 16 hybrid slots for extension modules, a PXIe system timing slot and a host controller slot. The modules are interconnected using a high-throughput PXIe backbone, as shown in Figure 28.

Figure 28 shows the hardware configuration of the DSP unit used during demonstration, including the extension modules and their interconnection. Installed in the chassis, there is a host controller (HC), ten programmable digital logic modules (FPGA), a signal generator (GEN), and a timing module (TM). Three of the DSP modules are connected to two data converters (ADC and DAC) and to an input/output module (I/O), respectively. The complete component listing of the DSP unit is given in Table 5. A photograph of the DSP unit configured in a *baseband-only* setup is shown in Figure 29.

The HC is an Intel x86-based, real-time embedded controller used to control the various chassis modules, and to interface them with a PC running *NI LabVIEW*. LabVIEW [182] is a graphical, data-driven programming environment with all chassis hardware directly accessible. The instrumentation PC is used to build the program onto the HC, and to send non-real-time configuration data to the HC through an

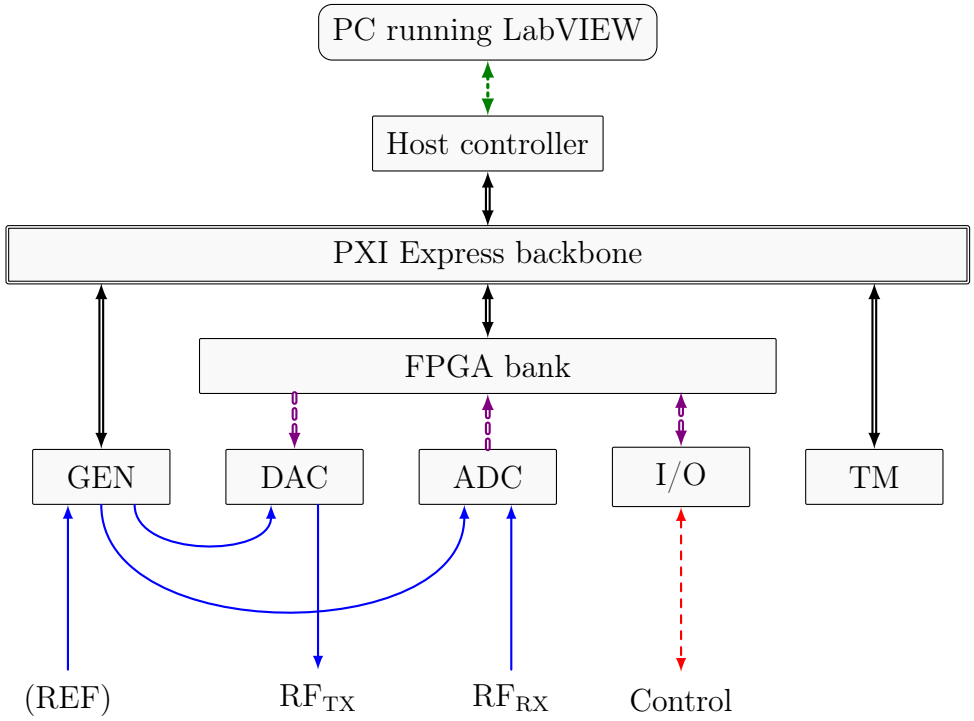


Figure 28: National Instruments DSP platform.

Table 5: Modules installed in the DSP unit.

Symbol	Model	Description	Slot	Ref.
DSP	NI PXIe-1085	PXI Express chassis	–	[173]
HC	NI PXIe-8135 RT	host controller	1	[174]
I/O	NI 6581	digital input/output	2	[175]
GEN	NI PXI-5652	signal generator	3, 9	[176]
DAC	AT-1212	digital-to-analog conv.	5	[177]
ADC	NI 5771R	analog-to-digital conv.	6	[178]
TM	NI PXIe-6674T	timing module	(10)	[179]
FPGA	NI PXIe-7966R	programmable logic	2, 5	[180]
FPGA	NI PXIe-7975R	programmable logic	6, 11 – 16, 18	[181]

Ethernet connection. The PC simply runs the graphical user interface or the *front panel*, as it is referred to within LabVIEW, relaying information to and from the chassis. While the HC runs LabVIEW Real-Time code, the information exchanged is not timing-critical.

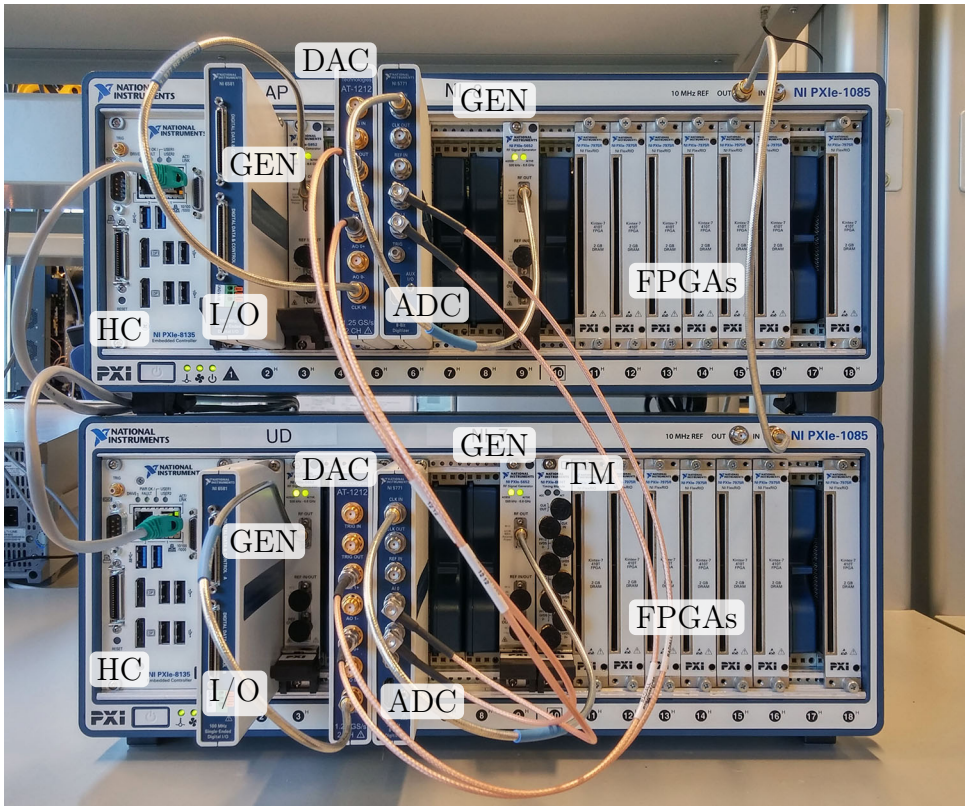


Figure 29: The DSP unit in a baseband-only configuration.

All timing-critical functionality is carried out on multiple field-programmable gate array (FPGA) digital logic modules. Using 1+1, 7+1 and 1 modules for transmit encoding, receive decoding and digital I/O, respectively, provides high degree of parallelism and throughput via peer-streaming. The FPGAs are responsible not only for all digital baseband processing but also for resource and link management. The latter involves controlling all external modules as shown in Figure 27, including those used in beam alignment.

The FPGA modules interact with the HC and each other through the PXI Express backbone. The data converters and the I/O module are directly connected to the FPGA cards. The digital I/O module interfaces with the various external blocks using multiple control signals, most of which are real-time. The data converters interface with the radio frequency subsystem, and are externally connected to the IF module: the DAC to the transmit branch (TX), and the ADC to the receive branch (RX). The signal generator is used to provide a high-precision clock signal to the data converters, possibly using an external reference oscillator (REF).

5 Network expansion

Together with the next section, this section discusses the automatic alignment functionality required in mmW APs deployed as a part of a self-organizing heterogeneous network (SOHN). Two distinct tasks may be identified: automatic network expansion and active misalignment compensation. The former refers to creating new links while the latter refers to maintaining and optimizing the currently used link. While these tasks are very different in nature, they are equally important for successful deployment of the envisioned network architecture. Given the number of network nodes, all costs should be minimized without degrading network performance [10], [11], [16], [29], [35], [36]. Intelligent beam alignment is a key feature in achieving these goals.

This section concerns the network expansion functionality while misalignment compensation is the topic of Section 6. These two sections also include author's original contributions. Our discussion on building the network is outlined as follows. In the first subsection, an overview of the alignment task is given. This is followed by discussion on the events taking place when a new node is activated in Section 5.2. The search functionality and co-operative pairing are discussed in Section 5.3. The last subsection compares the different approaches and their practical implications. Even though the discussion on these topics is mostly general in nature, there exists a clear connection to the target system described in Section 4.

5.1 Task overview

Network expansion and new link establishment are rare occurrences, and have mostly non-real-time implications [10], [16]. In most cases it makes little difference if the network is updated after some minutes rather than a few seconds after the new device has been powered up. However, timely and agile deployment are still required in situations where the network is temporarily extended. To name a few examples, these situations include public events and other mass gatherings as well as emergencies. Such situations are also considered within the METIS project under *Test Case 9: Open Air Festival* and *Test Case 10: Emergency communications* [8], [9]. In the latter, the network should be up and running within mere ten seconds.

These situations are, however, special occasions that warrant special hardware and may override typical best-effort operation. In most situations temporal requirements are relaxed but the expansion should not interfere with the existing service. While new *inoperational* nodes carry no traffic and can thus devote all of its resources for setting up new links, the existing network should concentrate on service continuity. The effort required from the *operational* nodes should be minimized.

Wired networks are expanded by installing another network switch, i.e., a router [162]–[164]. After the router has been initially configured and activated, it starts sending neighbor discovery queries to its immediate neighbors. Based on these

queries, other devices in the network are notified of new routes. They in turn share their connectivity information so that the new node may build a routing table.

While SOHNs are expanded in an analogous manner, the wireless nature changes the process somewhat. Instead of routing packets to an interface as soon as they are processed, they are sent to a certain direction at a certain time. Regardless, the first task in new node deployment is to preconfigure the networking device [84], [99], [100]. This initial configuration includes settings regarding default operation and network discovery. Most importantly, it identifies the operator, i.e., the owner, and provides all necessary functionality required to join the existing network.

To be used as the initial configuration for all nodes within certain geographical area, this configuration is not site-specific, i.e., local, but semi-global instead. The node may be configured either by the equipment manufacturer or by the operator prior to deployment. The site-specific operational configuration is then automatically downloaded after the node joins the network. It is later automatically updated to reflect any changes in the system.

Forming new links comes right down to neighbor discovery, i.e., mapping and connecting to existing network nodes. There are, however, a number of challenges. First, given the high directivity of the used antennas, proper alignment is required before a connection is possible. Second, both ends of the link need to be aligned using only limited means of feedback. Third, new nodes have not initially been synchronized. These features impose constraints that complicate the system considerably.

Prior to discussing constraints imposed on the system, it is necessary to discuss the beam alignment process in detail. First, all relay operations are assumed to be put on hold for the duration of alignment process. This is done for two reasons. First, a TDD node may only transmit or receive at any one time. Second, no transfer can be guaranteed using an unverified beam.

It should also be noted that the nearby environment may include networking equipment used not only by the intended operator but also by the competition. Furthermore, the radio environment might include incompatible devices such as user equipment (UE) and interferers. Moreover, as discussed in Section 3, a single source may also be visible in more than one direction due to multipath propagation. Additionally, multiple different sources may be seen in the same direction due to both finite spatial resolution and common scatterers [89], [90].

It is thus vital to differentiate different sources and to limit connections to valid targets. The process itself comprises two parts; there is a search phase and an authentication phase. During the search, the node scans the environment for other devices in similar fashion as a radar. Upon making contact, the nodes undergo a handshaking process and the authentication phase begins. During the authentication phase, nodes first exchange device identification, followed by various configuration information such as those related to synchronization, scheduling and network topology.

Missing synchronization imposes some constraints on the network expansion functionality. This is especially consequential in terms of the target system since

time-division is used for both duplexing and multiplexing. In essence, there are four solutions. First, if the node is deployed outdoors and is equipped with a suitable hardware, synchronization may be based on *Global Naviation Satellite System* (GNSS) clocks [83], [84], [99]. Another possibility is to equip the AP with a low-frequency radio and obtain synchronization using a network-distributed timing protocol such as the IEEE 1588v2 *Precision Timing Protocol* [183] (PTP). Third, the installation procedure involves a step for synchronization, for instance by using a specialized configuration-synchronization dongle.

It is also possible to design the alignment algorithm in such a manner that the endpoints need not be in synchrony [69], [184], [185]. By proper procedure and purposeful signaling, the new node becomes synchronized during the expansion process. To minimize the effect of missing synchronization, the search should be carried out by the operational node to avoid unnecessary pollution of the airwaves. More importantly in TDD systems, the transmitting node must preempt itself while the receiving node may react to external events like receiving a training message [83], [84], [88], [89], [99], [167], [168]. This pilot should be constructed with great care as it has to contain all the necessary information for the floating node to respond, and to account for both inter-symbol and inter-message spacing.

This is the starting point of our discussion on network expansion: minimal node hardware and functionality. While asynchronous beam training is possible, it is also both very conservative and extremely time consuming. Away from normal operation, it increases latencies and can thus decrease the quality of service (QoS) considerably. The rest of this section investigates different means of improving the alignment process while weighing their benefits and costs.

5.2 Node discovery

The fundamental question in mmW neighbor discovery is how to align both ends of the link at the same time. Throughout the discussion, *exhaustive search* [186] is used as the alignment algorithm. Exhaustive search is motivated by its simplicity and proven robustness [57]. While there are certainly more advanced alignment solutions, numerous algorithms are based on the same idea as is discussed in Section 7. The approach is a valid alternative for one-time events, and highlights different phenomena clearly. Furthermore, it is also one of the few algorithms compatible with all of the different beam steering implementations discussed in Section 2.

In exhaustive search, all possible beam pairs are tested blindly in a trial-and-error fashion, and the best beam is chosen for further communication. If there are N and M ($N, M \in \mathbb{N}^+$) possible beams at the two ends of the link, there are NM possible beam pair combinations in total. Assuming similar configuration at both ends of the link, the (asymptotic) complexity is quadratic, i.e., $\mathcal{O}(N^2)$.

This quadratic relation is not a problem for the inoperational node since it can devote all of its resources for the search. As a result, these resources are not as valuable as those in an operational node. The node can simply run the algorithm

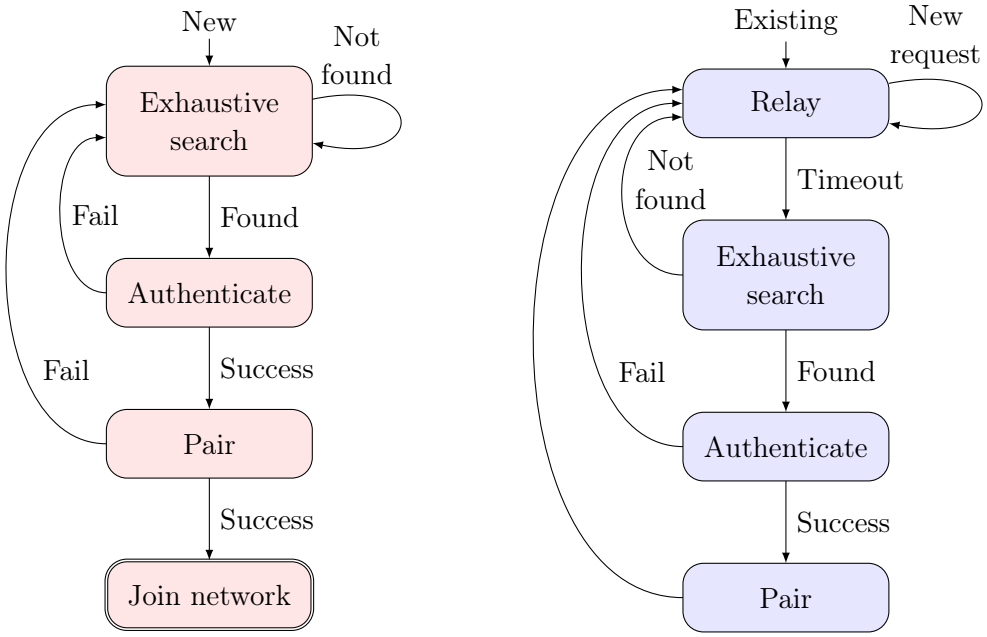


Figure 30: Automata describing the initial, blind approach for both the new (left) and existing (right) nodes.

until it joins the network. The existing node must conserve its resources, perhaps running the algorithm occasionally to see if there are any unpaired nodes. Automata describing this initial arrangement are given in Figure 30 for both the new and existing nodes.

The question remains how often the existing nodes should run the algorithm. When the new nodes are added, a shorter scan period is helpful. However, after the network has been built, it may run months – even years – without any changes. The scan delay is significant, and needless scanning must be avoided. While the needless scanning could be reduced by updating the network configuration after the network is built, it requires intervention and does not solve the issue completely.

Instead, the impact of new node establishment be minimized if the expansion algorithm is run only *on demand*. Moreover, the network is much more responsive – a highly desirable quality for temporary extensions. As shown in Figure 31, the network should be notified when a new node is ready to join the network. This would trigger the network to schedule a suitable network node to pair with the new device.

There are two ways to organize such functionality. The installer may trigger the search function manually after the device has been booted up, or the process can be automated if the AP is equipped with a secondary low-frequency radio. In the macro-assisted approach, the node would first connect to the macro BS and declare its existence. While additional radio increases the bill of materials, it can help to reduce the labor costs and provides additional features. For instance, the radio could be used for all low-rate communication such as feedback, synchronization and device configuration.

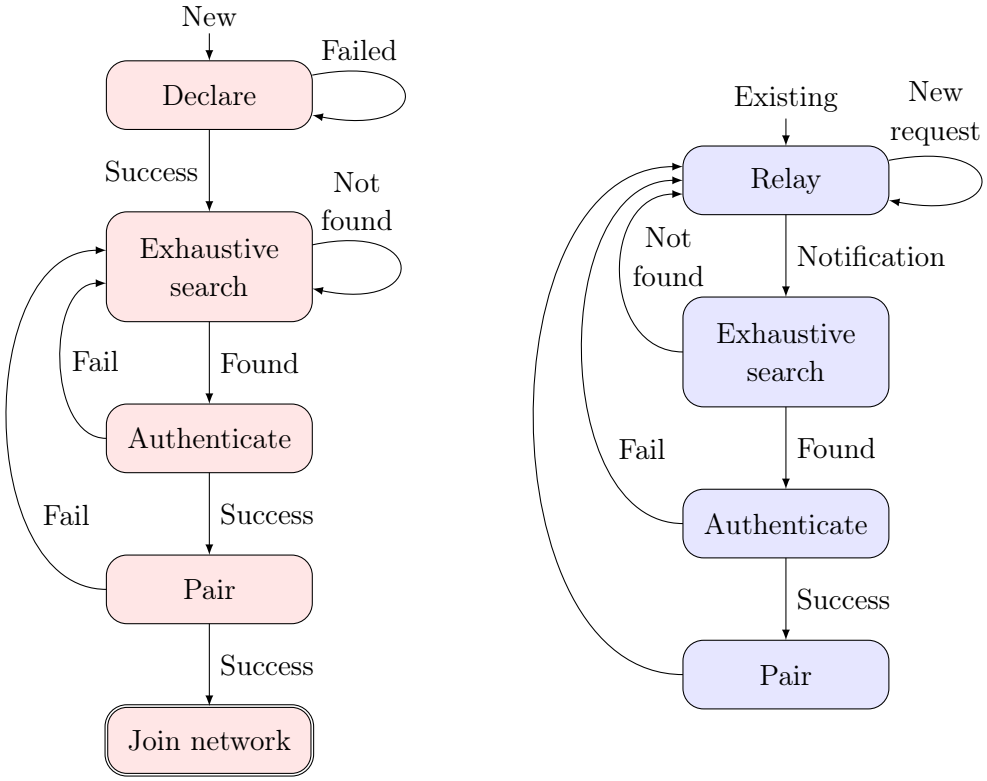


Figure 31: Automata describing the on-demand approach.

On-demand joining eliminates all needless scanning but does not reduce the complexity of a single search. The entire process could be further improved if the alignment process is made more efficient. A possible solution is to narrow down the search to a certain area. For this, the AP would need position information of the network devices within range – including its own.

After installation, the new node would first acquire an estimate of its location. Upon connecting to the macro network, the node would declare its location. This information would be conveyed to the nearby nodes available for pairing. The locations of these nodes would in turn be shared with the new node. Now, the nodes can search only if needed and target a certain area. Figure 32 portrays this arrangement.

In location estimation, global navigation satellite systems could be used [83], [84], [99]. With stationary targets and long integration times, GNSSs can resolve *absolute location* with reasonable accuracy. In addition to providing location information, GNSS clocks can be used for device synchronization. These systems, however, are available only where there is good enough satellite coverage. Moreover, they do not provide any information regarding the *orientation* of the device. Pairing requires *relative locations* obtained by combining absolute locations with orientation information. For this purpose, an orientation sensor (OS) could be used [187], [188].

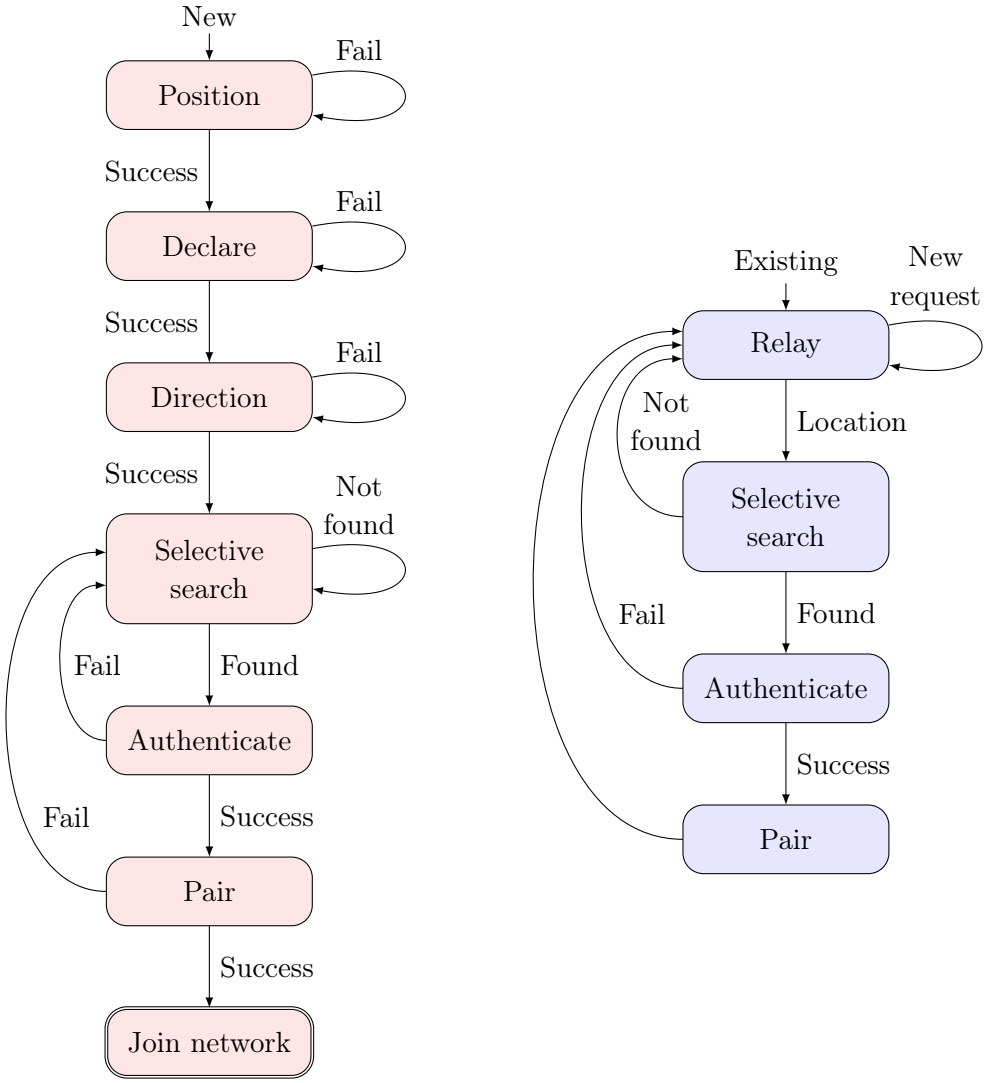


Figure 32: Automata portraying the macro-assisted approach with location information.

However, not only does a GNSS receiver increase the bill of materials, satellite-based positioning cannot be used indoors or in other places with no clear view of the sky. In these cases, one solution could be to extend the installer-assisted approach to also account for positioning. In addition to having a complete description of the existing network including device models, configurations and install locations, the network expansion is planned beforehand. While the planning focuses on identifying areas where extra capacity is needed rather than individual link planning, a complete installation itinerary will regardless be available.

The install location and orientation is predetermined to an acceptable accuracy without any hardware. This position information regarding all nearby devices could be uploaded into the device before installation. Alternatively, it could be downloaded after power-up using the macro link. At the same time, the information is also made

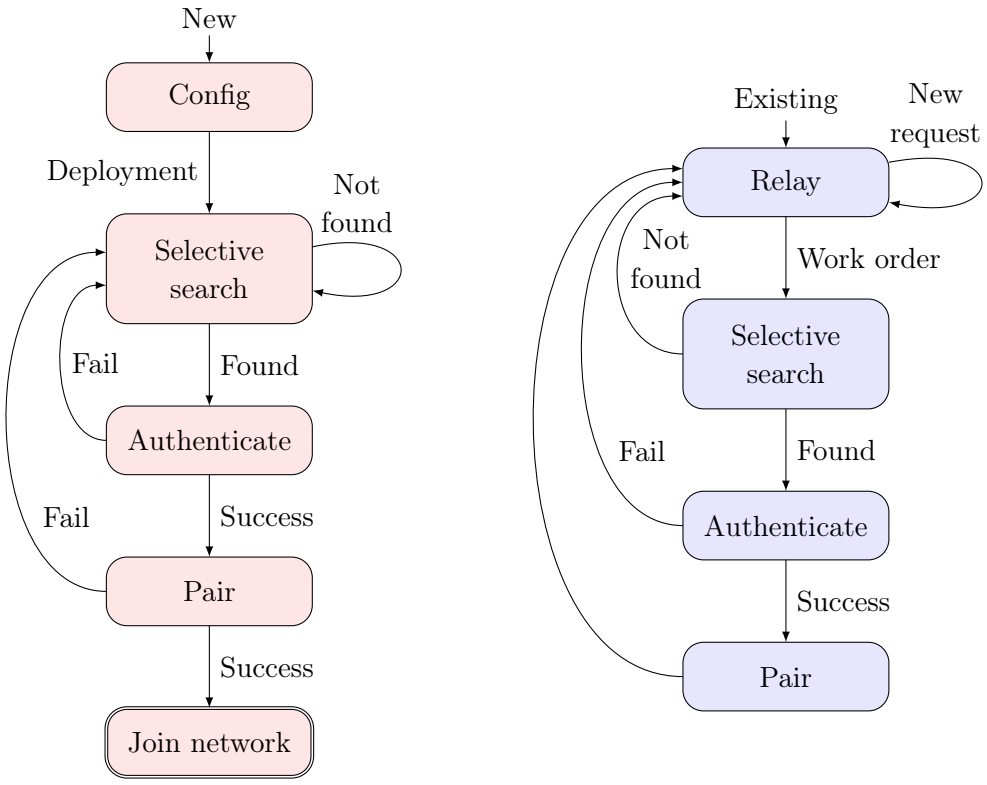


Figure 33: Automata describing the human-assisted approach utilizing preconfiguration and work orders.

available to nodes in the existing network. Upon completing the work order, the node will be added to the established network. This would signal for (one of) the existing nodes to begin their search, concentrating on the location of the new node. The corresponding automata are given in Figure 33.

While position information could indeed prove useful, it does not guarantee the device is actually seen in said direction. The line-of-sight path may be blocked, and the device is seen elsewhere or not at all. Moreover, the same device can be simultaneously visible in multiple directions due to multipath propagation. In a dynamic environment, the location or locations the node is seen in can even vary over time. A previously visible node might also become unreachable all of a sudden.

Linked closely with the network organization and management functionality, the beam pairing algorithm should also take these circumstances into account [29], [36]. Considering node availability, node failures may result either from external or internal issues. The failure may also affect the whole node and all of its links, or just the links in certain sectors. Four different cases may be identified as follows.

First, the node either crashes or is damaged, and becomes completely inoperational without any prior warning. The node is removed from the network, and the network recalculates its routing tables [162]–[164]. The dropped node should automatically recover by attempting to rejoin the network. If this is not possible, maintenance is

required. Thanks to high degree of automation, the node can simply be replaced by another with minimal effort.

Second, there is either a hardware or software issue limited to certain sectors. The node remains otherwise functional and connected to the network, and may continue serving requests with limited coverage. The node will provide the network with information about its state, and the network controller makes necessary changes to minimize disruption in service. If complete functionality cannot automatically be restored, operator intervention is required.

Third, the node becomes completely unavailable because of external reasons such as blockages in the signal path. In this case, none of the previously available links may be used anymore. The node is dropped from the network and should later attempt to rejoin the network. Fourth, some of the existing links become unavailable because of blockages but the node remains part of the network. Depending on the cause, the operator may have only very limited means of restoring full functionality.

Should a node be completely dropped, the situation is similar to the case where a new node is deployed. The only difference is that both the new node and the existing nodes alike already know where to concentrate search effort. The join operation is triggered when the node becomes unreached. The disruption may last for a while, meaning the search algorithm may need to run a couple of times. Hence, the discovery algorithm should run periodically until the connectivity is restored. To reduce the burden on connected nodes, the algorithm could implement an increasing backoff as is done in carrier sense multiple access (CSMA) systems [162]–[164].

When the node in question remains active and part of the network, it cannot be considered a new node. However, the aforementioned principles apply regardless. At both ends of the link, periodic and possibly backoffed search operation is undertaken to restore the link. In this case, neither endpoint can devote all resources to the search but have access to all necessary information. The pairing operation is now carried out concurrently with regular operation. This is also the case when a newly-installed node connects to more than one node in the mesh.

In addition to resource constraints, there are also timing constraints. While it is acceptable for new node deployment to take some time, the network should recover from outages as soon as possible. On the event that the main path is blocked, the network could also be restored using one of the multipaths. One of the alternate paths could be used to restore connectivity promptly if they are recorded before the primary path is lost. In this case, the restoration process would first try the primary path, followed by alternate paths based on the observed link quality.

Alternate paths are useful not only for robustness but also for improved throughput if both ends of the link support SU-MIMO techniques. These paths are discovered by mapping the radio environment after the node has connected to the network. The AP can map everything blindly, or coordinate the discovery only to form paths between nodes in the same network.

Table 6: An example of a radio environment mapping table.

#	**Owner**	**ID**	**Quality**	**Angle**	**TTL**
1	A	1	7	+180 °	452
2	B	3	3	+30 °	165
3	A	5	4	−90 °	407
4	A	1	5	+120 °	393
5	C	6	8	+30 °	97
6	A	1	4	−150 °	71
7	C	8	4	−30 °	299
8	×	×	6	±0 °	575
9	B	3	2	+120 °	204

The findings of the environment mapping function could be stored in a table similar to Table 6. In this example, the environment shown in Figure 34 is mapped blindly, i.e., completely. The table lists all the sources and paths currently identified by the node, one path per line. For each path, path index (#), owner, unique nodal identifier (ID), signal strength or quality metric, direction, and time to live (TTL) are listed. In addition to these, routing tables and link configuration information need also be stored.

Endpoints are identified using the operator and device identifier fields. A single endpoint may show up in multiple directions and vice versa. Multiple endpoints may be seen in the same approximate direction due to both finite angular resolution and common scatters. In the example, the angular resolution of A0 is highlighted with small ticks. Nodes A1 and B3 are simultaneously seen in three and two directions, respectively. Nodes B3 and C6 as well as the multipath components from nodes A1 and B3 are seen in the same approximate direction. Even though the example includes only the azimuth angle, both the elevation and azimuth angles need to be considered in practice.

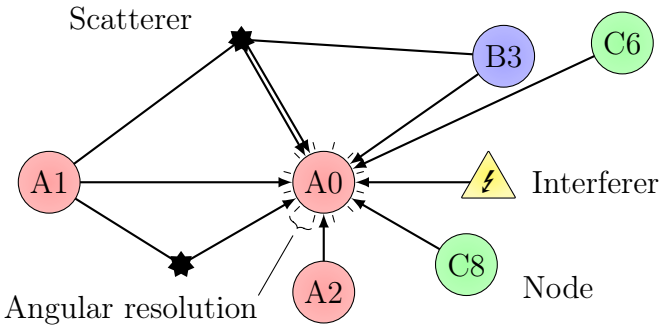


Figure 34: The radio environment in the example given in Table 6.

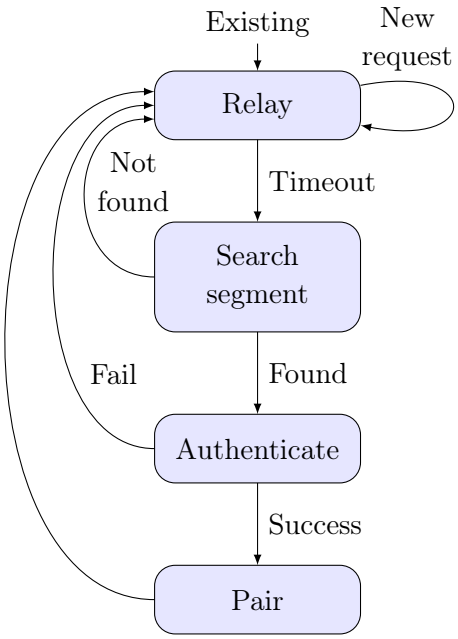


Figure 35: An automaton describing the segment-wise search used in network completion.

An interferer is shown with a cross (×) in both operator and device identifier fields. The link quality can be determined using received signal strength indicator (RSSI), signal-to-noise (SNR) ratio, or from received feedback. In the example, an arbitrary 0 ... 10 quality metric is used. Given the dynamic environment and multipath effect, the paths may not only change over time but coincide with other sources. For this reason, a TTL field or a time stamp is used to automatically remove any invalid paths.

Under normal operating conditions, the mapping function is not a time-critical task. The initial link is formed quickly, followed by a gradual increase of functionality. The network can be completed during off-hours when there are excess resources. To further lessen the impact on regular operation, the search area can be divided into small segments. This would decrease the length of the disruptive search since only a single segment is scanned at a time. The entire search area is covered over multiple runs, meaning the AP need not sacrifice performance or functionality. The corresponding automaton is shown in Figure 35.

5.3 Implementing the pairing function

The previous discussion only considered *what* to scan: everything or only a certain area, *when* to scan: periodically or only on-demand, and *why* to scan: not only extend the network but also to improve both availability and performance. It was not, however, considered *how* the nodes actually align their beams. In the following, we discuss how this is implemented in practice.

There are two different purposes for scanning the environment, and the search operation is carried out in different circumstances. To expand the network, initial links are formed with previously disconnected devices. Without macro assistance, the existing nodes cannot exchange any information any floating nodes. Information exchange is, however, possible when existing nodes seek to form alternate paths.

In any case, no information is exchanged using the mmW link during the search until both TX and RX beams have been aligned. If there is no secondary control connection, the search is carried out independently at both ends of the link. Each endpoint has only control over how different beams – not *pairs* – are tested. Moreover, the two endpoints cannot operate identically for the pairing to be successful. There are two reasons for this. First, uplink and downlink must be separated somehow, either in time (TDD) or in frequency (FDD) [83], [84], [88], [89], [99], [167], [168]. Second, using similar scan patterns at both ends would not cover the entire search area.

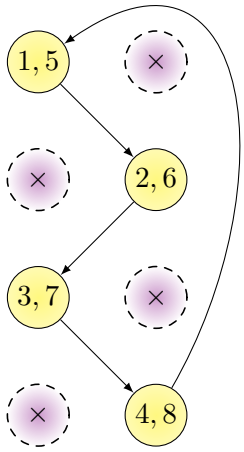
To elaborate the latter point, consider the following scenario. Assuming distinct, enumerated beams, i.e., a codebook, at both ends, the search can be considered searching for a particular cell in a two-dimensional matrix [52], [72], [120]. The dimensions of the matrix are N and M ($N, M \in \mathbb{N}^+$), corresponding to the number of possible beam directions at two ends of the link. During the discussion, we study an example arrangement where $N = 4$ and $M = 2$. The total number of possible beam pairs is thus $NM = 8$.

Consider the case where both endpoints operate identically, mirroring each other. Both ends of the link repeatedly scan through possible beam directions independently in the same order and at the same rate. The endpoints thus advance to the next option always at the same time. The scan sequences are too similar, and start to overlap very quickly. As a result, only 50 % of the search area is covered in the example. This arrangement is shown in Figure 36 where both the pseudocode and corresponding scan pattern are shown.

The given code is not the endpoint code as such but instead models the overall situation taking both ends into account. Furthermore, the code has been written with an emphasis on clarity rather than on performance. The code for a single endpoint is similar but one-dimensional. Even though the code given in Figure 36 is parametric, the scan duration is assumed to fit exactly NM different training messages. The comments in red show the results of different operations in our example case.

The scan functions as follows. After initializing the beam option variables N and M, the scan sequences at both ends of the link $\mathcal{I}$ and $\mathcal{J}$, respectively, are generated. The endpoints generate their own sequences and carry out beam tests accordingly. Now both ends of the link use effectively the same sequence generator to repeatedly test each position in order. The $\mathrm{rep}(\mathcal{S}, N)$ repeats sequence $\mathcal{S}$ N ($N \in \mathbb{N}^+$) times:

$$\mathrm{rep}(\{\underbrace{n_1, n_2, \ldots, n_K}_{K \text{ elements}}\}, N) = \{\underbrace{n_1, n_2, \ldots, n_K, n_1, n_2, \ldots, n_K}_{NK \text{ elements}}\}. \tag{19}$$



1: $N \leftarrow$ beam options at one end ▷ 4
2: $M \leftarrow$ beam options at the other end ▷ 2
3: $\mathcal{I} \leftarrow \text{rep}(\{1, 2, \ldots, N\}, M)$ ▷ 1, 2, 3, 4, 1, 2, 3, 4
4: $\mathcal{J} \leftarrow \text{rep}(\{1, 2, \ldots, M\}, N)$ ▷ 1, 2, 1, 2, 1, 2, 1, 2
5: **for all** $i \leftarrow \mathcal{I}$, $j \leftarrow \mathcal{J}$ **do** ▷ (1, 1); (2, 2); . . .
6: pair(i, j)
7: **end for**

Figure 36: Identical behavior at both ends of the link; pseudocode and the resulting scan sequence achieving half-coverage.

With scan sequences

$$\mathcal{I} = \{1, 2, 3, 4, 1, 2, 3, 4\} \text{ and}$$
$$\mathcal{J} = \{1, 2, 1, 2, 1, 2, 1, 2\},$$

only positions $(1, 1)$; $(2, 2)$; $(3, 1)$ and $(4, 2)$ are tested. Only half the possible beam pairs are tested, and that half is tested twice. In the general case, the number of distinct pairs may be found as the least common multiple (LCM) of the two dimensions N and M. Coverage C may thus be found as

$$C = \frac{\text{lcm}(N, M)}{NM}. \tag{20}$$

Full coverage is achieved only when $\text{lcm}(N, M) = NM$. Since $\text{lcm}(4, 2) = 4$ and $4 \cdot 2 = 8$, only a half of the search area is covered.

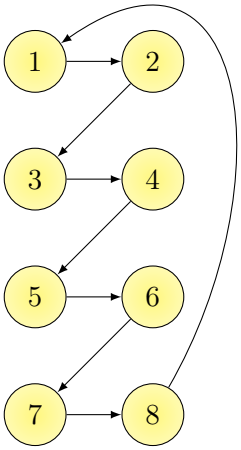
Equation (20) is easily justified by investigating scan sequence overlap. The pairing sequence is obtained by picking elements in order from sequences of the format shown in Equation (19). We are interested in the number of *distinct*, i.e., unique, *ordered* pairs ($(N, M) \neq (M, N)$) formed with this logic. The in-order $\mathbb{N}^+$ test sequences align perfectly every $\text{lcm}(N, M)$ elements, and the pairing sequence starts repeating itself.

This idea is elaborated in Table 7 where the alignment of three $\mathbb{N}^+$ sequences of different lengths are shown. Test sequences are formed by picking elements in order from two rows. For original sequences of 2 and 4 elements each, the first $\text{lcm}(2, 4) = 4$ ordered pairs are unique. Similarly for the other combinations, $\text{lcm}(2, 3) = 6$ and $\text{lcm}(3, 4) = 12$. In the table, the first ordered pair that is repeated is highlighted using different colors for each combination.

The resulting search pattern is given on the left in Figure 36. Indices i and j ($i, j \in \mathbb{N}^+$) refer to current row and column, respectively. The circles represent

Table 7: Sequence overlapping with different sequence lengths.

#	1	2	3	4	5	6	7	8	9	10	11	12	13
$\{1,2,3,4\}$	1	2	3	4	1	2	3	4	1	2	3	4	1
$\{1,2,3\}$	1	2	3	1	2	3	1	2	3	1	2	3	1
$\{1,2\}$	1	2	1	2	1	2	1	2	1	2	1	2	1



1: $N \leftarrow$ beam options at one end ▷ 4
2: $M \leftarrow$ beam options at the other end ▷ 2
3: $\mathcal{I} \leftarrow \text{sort}(\text{rep}(\{1,2,\ldots,N\}, M))$ ▷ 1, 1, 2, 2, 3, 3, 4, 4
4: $\mathcal{J} \leftarrow \text{rep}(\{1,2,\ldots,M\}, N)$ ▷ 1, 2, 1, 2, 1, 2, 1, 2
5: **for all** $i \leftarrow \mathcal{I}$, $j \leftarrow \mathcal{J}$ **do** ▷ (1, 1); (1, 2); . . .
6: pair(i, j)
7: **end for**

Figure 37: Fast-slow approach; pseudocode and the resulting scan sequence achieving full coverage.

distinct beam pairs. Tested directions are drawn using solid circles and numbered in the search order while unchecked pairs are marked with a cross ($\times$) and highlighted with a dashed line. The arrows correspond to transitions between the positions.

Mirrored behavior is thus not a viable solution, and some distinction between the two endpoints is required. The situation is somewhat similar to CDMA systems where different users are not separated in time, in frequency or spatially but are distinguished by unique codes [83], [84], [163]. The overlaid signals may be recovered if the codes are *orthogonal* to one another. Using a similar idea, there are two solutions for improving the test sequences to achieve better coverage.

One approach is to assign the endpoints different roles and different sequence generators designed to work together. A straightforward solution is to consider the situation as a nested loop with the two test sequences as the two loop indices. Referred to as the *fast-slow* division, the nested-loop approach is outlined in Figure 37. Full coverage is achieved when the *slow* endpoint switches to another beam only after the *fast* node has tested all of its beams, and full coverage is achieved. The code given in Figure 37 is identical to that shown in Figure 36 except sequence $\mathcal{I}$ is now sorted on line 3.

The other approach is to resort to random number generation (RNG) as shown in Figure 38. In the example, the test sequences are generated using a *shuffle-repeat* function. The shuffle$(\mathcal{S}, N)$ function returns a sequence obtained by concatenating

N independently shuffled copies of the sequence $\mathcal{S}$ similarly to the following example:

$$\text{shuffle}(\{\underbrace{n_1, n_2, \ldots, n_K}_{K \text{ elements}}\}, N) = \{\underbrace{n_2, n_K, \ldots, n_1, n_K, n_1, \ldots, n_2}_{NK \text{ elements}}\}. \tag{21}$$

Even though the endpoints generate their own sequences in an identical fashion, the random element inherent in the shuffling operation keeps the two sequences from mirroring each other constantly.

In the example, the two random-generated sequences are

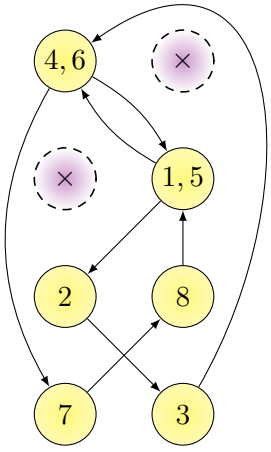
$$\mathcal{I} = \{2, 3, 4, 1, 2, 1, 4, 3\} \text{ and}$$
$$\mathcal{J} = \{2, 1, 2, 1, 2, 1, 1, 2\}$$

which correspond to the pairing sequence

$$\{(2,2);\ (3,1);\ (4,2);\ (1,1);\ (2,2);\ (1,1);\ (4,1);\ (3,2)\}.$$

Beam pairs $(1,1)$ and $(2,2)$ are tested twice while pairs $(1,2)$ and $(2,1)$ are foregone. With these sequences, a coverage of 75 % is achieved. The randomized approach thus provides potential improvement over the case where the endpoints mirror each other. Random generation, however, does not guarantee full coverage or even any improvement over the mirroring approach.

It is important to note the sequences are generated through repeat-shuffling instead of pure random generation. Shuffling ensures both endpoints actually test all their own beam directions while pure randomization could keep both endpoints static. Sequences like $\{i, \ldots, i\}$ are thus forbidden. In the best case scenario, all NM beam pairs are tested. In the worst case, the beams are paired statically meaning i is always paired with the same j for all is. In the absolute worst case scenario, all is always map to the same js. This is possible when $\text{lcm}(N, M) = \max(N, M)$. Now



1: $N \leftarrow$ beam options at one end ▷ 4
2: $M \leftarrow$ beam options at the other end ▷ 2
3: $\mathcal{I} \leftarrow \text{shuffle}(\{1, 2, \ldots, N\}, M)$ ▷ 2, 3, 4, 1, 2, 1, 4, 3
4: $\mathcal{J} \leftarrow \text{shuffle}(\{1, 2, \ldots, M\}, N)$ ▷ 2, 1, 2, 1, 2, 1, 1, 2
5: **for all** $i \leftarrow \mathcal{I}$, $j \leftarrow \mathcal{J}$ **do** ▷ $(2,2); (3,1); \ldots$
6: pair(i, j)
7: **end for**

Figure 38: Randomized approach; pseudocode and one of the possible scan sequences. This time, 75 % coverage is achieved.

Table 8: Mean coverage rates achieved using the randomized approach averaged over 1000 instances.

%	2	4	8	16	32	64	128	256	512	1024
2	74.2	74.1	74.8	75.2	74.9	75.0	75.0	74.9	75.0	75.0
4		68.4	68.4	68.4	68.4	68.4	68.3	68.4	68.3	68.3
8			65.7	65.7	65.6	65.6	65.6	65.6	65.7	65.6
16				64.4	64.4	64.4	64.4	64.4	64.4	64.4
32					63.7	63.8	63.8	63.8	63.8	63.8
64						63.5	63.5	63.5	63.5	63.5
128							63.4	63.4	63.4	63.4
256								63.3	63.3	63.3
512									63.2	63.2
1024										63.2

only $\max(N, M)$ pairs are tested leading to a coverage of $\min(N, M)^{-1}$. In general, achieved coverage C is thus bound by

$$\min(N, M)^{-1} \leq C \leq 1. \tag{22}$$

Table 8 shows the mean coverage rates for some values of N and M. The rates are given in percent, and are averaged over 1000 iterations. The experiment was conducted using *MATLAB 2014b x64* [189] running on *Ubuntu 12 LTS* [190] with `randperm()` [191] as the random element. The source code is given in Appendix A. The values for N and M were chosen as power-of-twos since they correspond to the worst case scenario where $\text{lcm}(N, M) = \max(N, M)$. Regardless, more than 63 % of the search area is covered on average – even with unrealistically high number of beams. Using the mirrored approach, coverages would equal to $\min(N, M)^{-1}$ for all values of N and M.

While there are no predefined scan sequences, the process still needs to have clear rules and etiquette. Timing constrains are still enforced, and the endpoints must adhere to the same exact pairing schedule. As discussed earlier, the nodes cannot operate in an identical fashion but must have a clear division of labor through *roles*. The role determines the TX-RX division and messaging order; whether the node should be the one to send or to receive training messages at a certain point in time. The used scan sequence could also be determined by the role.

These roles may be agreed upon beforehand or randomized – possibly even for each attempt at pairing. A pure randomized solution would incorporate both random scanning sequence and role. Such an algorithm is simple since optimizing role assignment is not a concern. Furthermore, the same configuration may be used for

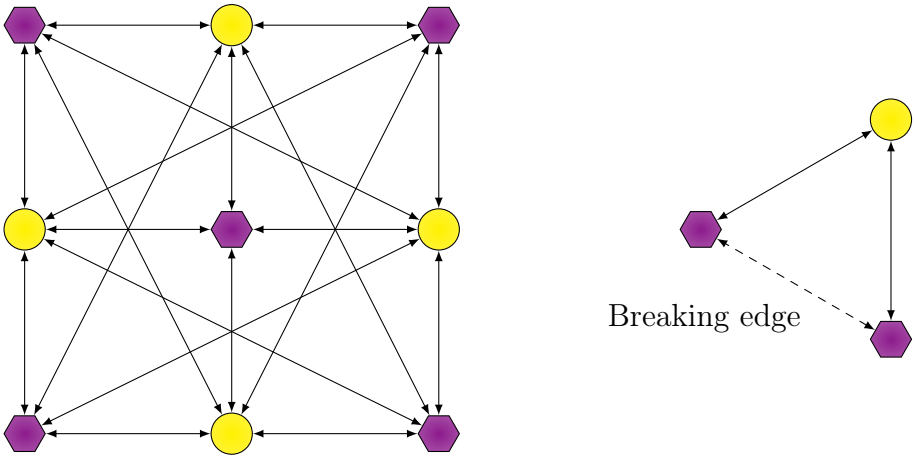


Figure 39: Examples of both bipartite (left) and non-bipartite (right) graphs.

each node in the entire network. This superior scalability is the greatest advantage of randomized approach over its deterministic counterparts.

The price paid is that the search area may not be completely covered. Even if the beams are perfectly aligned, the pairing might fail due to both ends being in receive or transmit modes at the same time. Not only does the randomized approach skip some beam pairs but also some pairing opportunities when both ends operate in the same mode. The expected pairing rate is thus half that shown in Table 8 [192]. Regardless, the algorithm would eventually be able to pair with all visible devices.

With deterministic approaches the key question is how the roles are assigned to different nodes. Four different approaches are considered assuming there are only two possible modes. While, in theory, it is indeed possible to have more than two different roles and scan patterns, it would not make much sense in practice. The links are always formed by pairing two nodes at a time – even if one or both nodes support P2MP functionality. In case of P2MP, however, there may be additional requirements in terms of simultaneous transmission and reception, as already discussed in Section 4.1.

The first approach to deterministic node assignment is to use *static* roles on a *global* scale. The network nodes are permanently divided into two groups, and only devices from different groups can pair with each other. All links may be formed only if the network is bipartite, i.e., two-colorable [193], [194]. An example of a bipartite graph is shown on the left in Figure 39.

One possible division is to separate nodes to active and inactive nodes based on whether they are part of the network or not. This way new nodes may always pair with any existing node within range. On the other hand, network nodes may only pair with inactive nodes. As a result, cycles are not possible, and the topology is limited to trees [193], [194]. Furthermore, the network builds solely on single paths. No alternate paths are used, meaning failure close to the root, i.e., gateway, would cripple the entire network.

Static assignment – let alone new-existing division – is not a viable solution since even simple networks may not be bipartite. A network as small as three nodes may not be two-colorable [193], [194]. This minimalistic example is shown on the right in Figure 39 where the edge breaking the bipartite property is highlighted. A better approach would involve *dynamic*, i.e., temporary roles, orchestrated either globally by the macro BS or locally by the links themselves.

Now the nodes switch roles either periodically or to react to some event such as the declaration of a new node. One possible solution is to incorporate the role in the network schedule discussed in Section 4.1. With dynamic roles there are no limits to network topology but a role assignment may forbid a node from pairing with some or any of its neighbors for the duration of the assignment. This restriction is, however, inherent in *any* role assignment since TDD essentially enforces two-coloring temporarily [193], [194].

The fourth approach is to randomize roles periodically while still following a deterministic scanning routine. The scheme is *node-centric*; the nodes assume roles independently of one another. There are no administrative overheads since no additional information is exchanged. The solution is easily scalable since no external management or agreement is needed. With expected pairing rate of 50 % [192], however, suboptimal division is more common than with the other two dynamic assignment schemes.

5.4 Practical considerations

Based on the previous discussion, deterministic means can easily be seen to outperform RNG-based approaches in terms of both scanning and pairing efficiency. Comparing the fast-slow and shuffle-repeat scan patterns, the former achieves full coverage while the latter can be expected to cover 63 % of the search area on average. In role assignment, randomized grouping can only be as good as a calculated division – never better.

The deterministic methods do indeed perform better but only if they are efficiently organized. This is a challenge especially if the network is built using dissimilar devices with different functionality and features. The nodes cannot run the same configuration but have to be properly managed to ensure compatibility. In addition to increasing system complexity and overheads, the management functionality is now also another source of errors, and possibly even a single point of failure could cripple the entire system.

Management overheads are regardless negligible in comparison to alignment overhead, and can be further reduced using secondary control connection. The dominant training overhead can also be reduced if the training is sped up by reducing the search area. In addition to physically limiting the search to certain directions, the *effective* search area can be decreased by increasing beamwidth.

Known as *floodlighting* in the context of radars [195], [196], the nodes would initially use a wide beam to acquire a coarse location which is further refined using

a narrow beam. Variable beamwidth and multi-level codebooks enable *hierarchical searching* similarly to a *K-ary search tree* [197]. Compared to linear search used thus far, the hierarchical approach provides significant speed-up by effectively reducing the search area logarithmically [52], [63], [76]. Reciprocally, the alignment accuracy could be improved by continuing the search further at the cost of increased effective search area and training time.

This idea behind hierarchical search is elaborated in Figure 40 showing a three-stage codebook for eight beams at both ends of the link. At the first stage, two top-level beams are needed to cover the entire scanning range. Pair 4 is measured to be most promising, and further search effort is limited to these directions. The search is refined in this manner until fine-grained beams are reached. Using binary search, fine alignment is obtained in just 12 training messages instead on the 64 required by linear search. However, the number of required feedback messages has tripled from one to three.

Increasing antenna beamwidth reduces gain, and thus decreases the SNR of the received signal [54], [55], [85]. Luckily, the information exchanged during pairing is low-rate, and the entire function is availability-centered rather than throughput-oriented. Consequently, the modulation scheme used in pairing can be simple and inefficient yet resilient toward bit errors. For instance, a low-order, highly-redundant phase-shift keying (PSK) scheme would be well-suited for the task [83], [118], [167]. The smaller the minimum SNR, the smaller the required gain. For instance, a 1/5-coded binary-PSK (BPSK) data stream can be recovered even if the noise and signal powers are comparable.

Given the exponential speed-up compared to the traditional exhaustive search, hierarchical codebooks are widely used in beam alignment [52], [63], [69], [76]. If supported by the hardware, the multi-level search is rather straightforward to implement. The link set-up protocol needs simply to include multiple alignment phases depending on codebook depth. However, the protocol must take into account that codebooks of different depth may be used at two ends of the link. Moreover, the

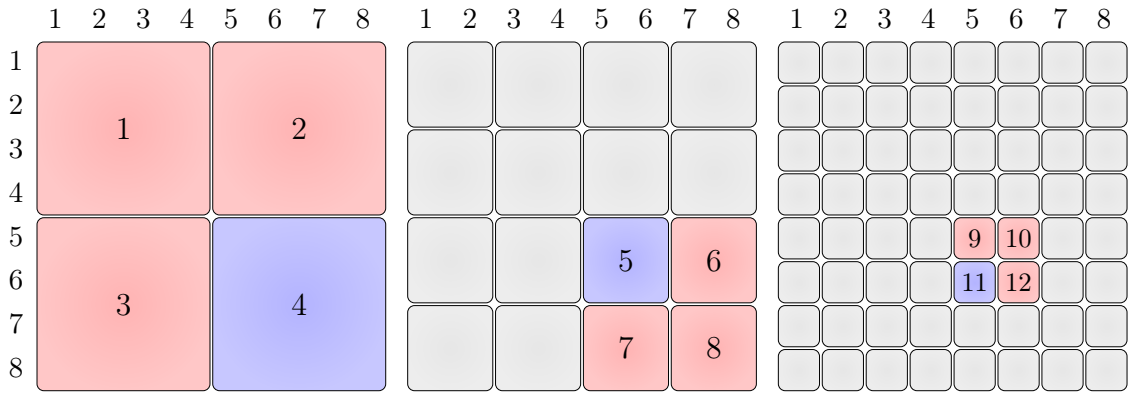


Figure 40: An example of a three-level binary search with 8×8 beam pairs.

link refinement procedure must be supported also by nodes that do no implement beamwidth control themselves.

While the previous presentation focused on backhauling, most of the principles are also valid for access links. On one hand, the two situations are fundamentally similar. On the other hand, though, they are completely different. First of all, APs and BSs are approximately stationary meaning the situation can be considered to be static. Furthermore, the number of backhaul nodes in a network is limited. It is also likely that most if not all visible devices are of interest, and are provided by a small number of vendors. Lastly, the AP placement is still somewhat planned even if rigorous network simulations are not carried out.

Considering user equipment, none of the aforementioned four points hold. There can be plenty of users that can move freely in the environment while holding their devices in different positions. The users may spend only mere seconds within the range of a single access point. In a crowd, the users and devices may also be very close to one another. The movement and arbitrary equipment positioning can also cause blockages in the primary path.

Between backhaul and access link, tot only is the environment more dynamic but also the devices are very different. Provided by multiple different manufacturers, the UE functionality will vary greatly from a device to the next. On a positive note, role assignment is trivial since access links are always formed between a UE and an AP – mmW ad-hoc links are not planned. The UE search space is also smaller, since the handset antennas are not nearly as directive as those used in APs [54]–[56], [62]. While this is because of the small antenna size, low gain facilitates alignment.

It is also worth noting that not all users have a subscription from the operator who owns the AP in question. Furthermore, only a small portion of the users within range require services from the high-throughput access node. The service is thus invoked only when the user requests a large data transfer. Moreover, disruption in mobile access is not as consequential as a disruption in backhaul since only a single user is affected.

For aforementioned reasons, actively mapping the radio environment may not be feasible. There are simply too many devices to cause unacceptable overheads. Moreover, the benefits would also be minimal as the accuracy and thus usefulness of the mapping would degrade after mere seconds. This is because the users are typically mobile, free to move outside the area covered by the AP at any time.

Pairing on demand could thus prove to be a more attractive solution. The pairing process would be initiated only when the user first requests the service or when the user moves to an area covered by another node. Both cases are macro-assisted, and could incorporate approximate location information. This way the AP only maps active connections, greatly reducing the workload at the cost of a small one-time initiation delay. The connections are, however, numerous and minimizing alignment delay is of vital importance. In these situations exhaustive search may not be as attractive a solution as the more advanced algorithms discussed in Section 7.

6 Misalignment compensation

This section considers how to maintain existing links in order to improve availability, throughput and reliability. Initial alignment is not enough as the radio devices will be installed in suboptimal locations using low-cost mounting structures [32], [34], [35], [44]. These installation sites are prone to small yet significant movement and vibration possibly steering narrow pencil beam antennas off their targets [13], [21], [30], [32], [44], [50]–[52]. Decreasing link quality, misalignment may affect throughput or even availability, and must be compensated in a dynamic fashion.

This section is organized as follows. First, the coordinate convention and associated terms are defined. Next, Section 6.1 investigates misalignment in theory by modeling an access point installed on top of a pole that is subject to twist and sway. This is followed by a short overview of measurements characterizing misalignment in practice in Section 6.2. Based on these findings, Section 6.3 introduces some possible approaches. Sections 6.1 and 6.3 include author's original contributions.

This thesis adopts the Cartesian coordinate convention used by Lee and Seshia [198] with the arrangements and terminology shown in Figure 41. The *longitudinal*, *vertical* and *lateral* axes are denoted with x, y and z, respectively. The rotation around these axes are referred to as *roll*, *yaw* and *pitch*, respectively. These are further denoted using α_x, β_y and γ_z, respectively. Because of this arrangement, the x, y and z axes are also referred to as roll, yaw and pitch axes, respectively. *Sway* refers to rotation in the roll-pitch plane $\alpha_x\gamma_z$ whereas *twist* corresponds to yaw β_y. *Lateral sway* corresponds to roll α_x while *longitudinal sway* is the same as pitch γ_z.

Interested in relative quantities rather than the absolute position, the coordinate system is used rather loosely. This is evident in two aspects, the first of which is the endpoint-specific coordinates. Section 6.1 always considers a situation where one of the endpoints is mounted in the origin $(0, 0, 0)$ facing toward the positive x axis. In

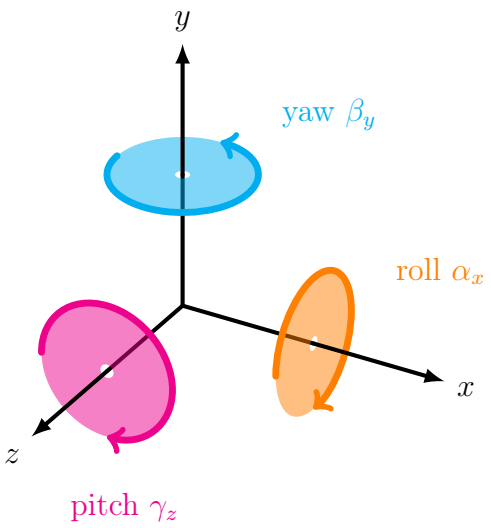


Figure 41: The coordinate convention used in this thesis.

the following, the x axis is assumed to point toward the other end of the link while y axis points toward the sky. Direction of the third orthogonal axis z can determined either by using the right-hand rule or as the vector cross product of the other two unit axis vectors: $\vec{z} = \vec{x} \times \vec{y}$.

The other aspect is the ambiguity of α_x, β_y and γ_z. A fixed reference or the sense is *not* explicitly defined for any angle. Instead, the presentation always refers to *relative angles* using a reference that is most useful in the specific situation. Typically this corresponds to the resting position. Furthermore, most cases are symmetric and the sense of rotation is unimportant. However, the relative distances are shown using a Δ. For instance, Δx is used to refer to distances along the x axis.

6.1 Modeling misalignment

Next, the effect of misalignment on link quality is studied by developing a simple geometrical model. The topic is approached through an example where an AP is mounted on top of a pole. In real life, this pole could be a lamp post, a traffic sign or a utility pole to name a few examples. Being a low-cost structure, the pole is prone to vibration caused by weather effects and other phenomena.

Resting, the pole points straight up toward the positive y axis. The pole is assumed to have a length h. The AP is faces toward another endpoint positioned on the positive x axis at a distance r and at the same height as shown in Figure 42. The devices are assumed to be identical and initially perfectly aligned.

For simplicity, the pole is assumed to be both perfectly straight and rigid while investigating the effects of both sway and twists. The AP and the pole are considered as a whole, assuming the structure pivots only from its bottom end. All other transformations such as bending and contraction are considered negligible. These assumptions preserve the orthogonality of roll, yaw and pitch, meaning their effects may be studied independently [199].

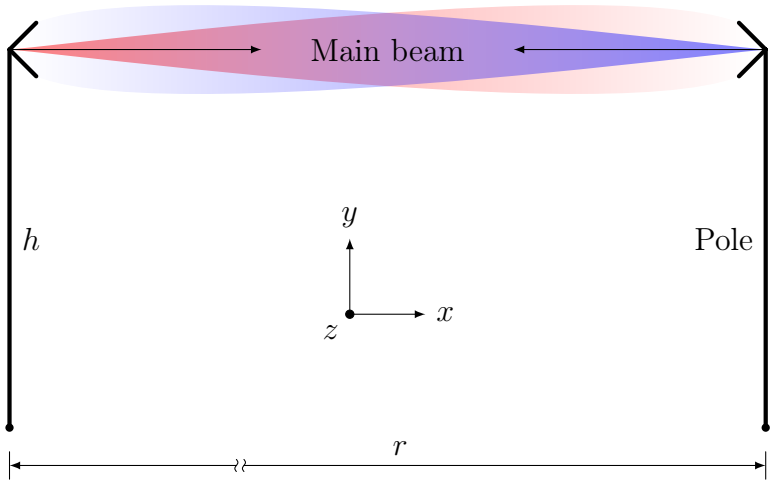


Figure 42: Modeled situation.

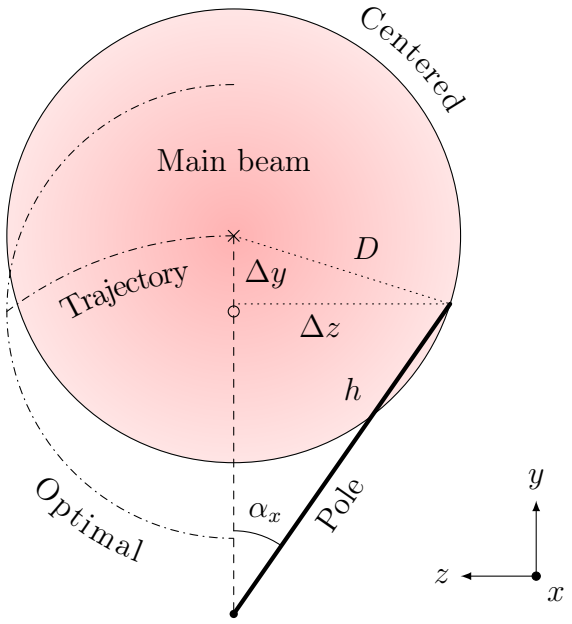


Figure 43: Misalignment caused by roll α_x.

The effects of lateral sway, i.e., roll α_x around the longitudinal x axis, are discussed first. Consider the situation pictured in Figure 43, showing a red disc that corresponds to the area illuminated by the other endpoint. Inside this disc, the received power is at least half of the maximum. The radius R of this disc may be found using the half-power beamwidth (HPBW) and basic trigonometry as

$$R(HPBW, r) = \sin\left(\frac{HPBW_\circ}{360^\circ}\pi\right) r. \tag{23}$$

Considering an antenna with $HPBW = 1.0^\circ$ placed $r = 100$ m away, the radius R is approximately equal to 87 cm. These values do not correspond to those shown in the figures. For the rest of the analysis, the radius R is assumed to be less than the pole height h, i.e., $R < h$.

The illumination disc is centered around the resting position of the endpoint, denoted with a cross ($\times$). If the antenna is placed anywhere inside this area, the link is assumed to be fully functional. Displacements D less than the radius R thus have negligible effect on link quality. A pole swaying from its root causes a circular displacement trajectory symmetric about the resting position on the y axis. The left half of this trajectory is shown in Figure 43.

While on this trajectory, the displacement may be found as a function of both the roll angle α_x (with respect to the positive y axis) and the pole height h. Assuming a perfectly circular beam, displacement $D(\alpha_x, h)$ is given in Equation (24):

$$D(\alpha_x, h) = \left\{ \Big[\underbrace{\sin\alpha_x\, h}_{=\Delta z} \Big]^2 + \Big[\underbrace{(\cos\alpha_x - 1)h}_{=\Delta y} \Big]^2 \right\}^{\frac{1}{2}}. \tag{24}$$

The orthogonal offset components Δy and Δz are also highlighted, dictating that the sense of α_x is counterclockwise.

Solving Equation (24) for α_x yields

$$\alpha_x = \pm \arccos\left(1 - \frac{D^2}{2h^2}\right) \tag{25}$$

assuming $D < h$. The roll angle causing the AP to leave the illuminated area is found by setting $D = R$. Assuming $R = 87$ cm as found in the previous example and a pole of height $h = 8.0$ m, a roll $|\alpha_x| > 6.25\,^\circ$ affects the quality of the link. Since the vertical offset Δy is always non-positive, i.e., downward, the roll limit may be maximized by directing the antenna slightly below the resting position.

This principle is shown on the left side of Figure 43. In the figure, the dashdotted half circle shows the optimal beam position below the target-centered beam. The center of the optimum beam is marked with a circle ($\circ$). Placed optimally, the maximum horizontal offset $|\Delta z|$ equals to R. Assuming $R < h$, this corresponds to a roll

$$\alpha_{x\max} = \pm \arcsin \frac{R}{h} \tag{26}$$

when the beam is positioned

$$\Delta y = h - \sqrt{h^2 - R^2} \tag{27}$$

below the AP. With the previously used values, the roll limit becomes $|\alpha_{x\max}| = 6.26\,^\circ$ if the beam is positioned optimally $\Delta y = 48$ mm below the target. With small beamwidths this optimization provides only a minimal benefit.

Based on the analysis and examples above, it is unlikely that roll would cause displacements with a significant impact on link quality. It would simply require too great a roll angle to be encountered under normal operating conditions, as is discussed in Section 6.2. In the former, though, perfect initial alignment was assumed. This is highly unlikely if not altogether impossible in practice. For instance, the alignment resolution of beam switched or codebooked antennas is typically on the order of HPBW.

Even with a beamforming antenna, perfect alignment would require phase shifters with an unrealistic accuracy. Instead, the antenna may often be positioned so that the connection is barely possible to begin with. In this situation, even a small roll could result in a noticeable misalignment – especially with longer poles. Moreover, the illuminated area is proportional to both beamwidth and link distance meaning the absolute worst setting involves a short hop with long poles and narrow antennas.

Roll causes both lateral and vertical displacement seen identically at both ends of the link. In other words, it does not make a difference which endpoint moves in the yz plane since both endpoints experience the same relative effect. Furthermore, this movement may be measured and compensated but only for the endpoint in question – a topic explored in depth in Section 6.3.

Roll degrades link quality simultaneously in two ways, the first of which is the displacement in the yz plane discussed above. This effect is possible only when the antenna is attached to a swaying structure with leverage. The antenna rotates, decreasing the performance of linearly polarized systems. Not only does rotation reduce gain but also introduce cross-polarized components. Effectively, this decreases the antenna cross-polarization discrimination (XPD) crucial for co-channel dual polarization (CCDP) systems.

XPD was already defined in Equation (17) during the discussion on channel modeling in Section 3.4. When discussing antennas, however, it can be rewritten using decibels as follows [86]

$$XPD_{\mathrm{dB}} = 10 \log_{10} \left(\frac{P_{\mathrm{C,\,max}}}{P_{\mathrm{X,\,max}}} \right) \mathrm{dB} = G_{\mathrm{C,\,max,\,dB}} - G_{\mathrm{X,\,max,\,dB}}. \tag{28}$$

Here $P_{\mathrm{C,\,max}}$ and $P_{\mathrm{X,\,max}}$ are maximum radiated powers of the co-polarized and cross-polarized components, respectively. Similarly, $G_{\mathrm{C,\,max,\,dB}}$ and $G_{\mathrm{X,\,max,\,dB}}$ are the maximum antenna gain in the two polarizations, respectively, given in decibels.

Assuming linear polarization, the two electric field components may be found by multiplying the original field magnitude by cosine and sine of the roll angle α_x, respectively. Defined for powers, the field magnitudes need to be squared: $P_{\mathrm{C}} = \cos^2 \alpha_x \, P$ and $P_{\mathrm{X}} = \sin^2 \alpha_x \, P$. The powers P are canceled in the fraction, and the XPD equals to one per tangent of roll α_x squared:

$$XPD_{\mathrm{dB}}(\alpha_x) = 10 \log_{10} \left(\frac{\cos^2 \alpha_x}{\sin^2 \alpha_x} \right) \mathrm{dB} = -20 \log_{10} |\tan \alpha_x| \ \mathrm{dB}. \tag{29}$$

The equation assumes the polarizations are perfectly aligned when both APs are at rest.

While a two-degree roll has virtually no effect on the magnitude of the co-polarized component (-0.0053 dB), a cross-polarized component with a magnitude of more than -30 dB is created. That is, even with a roll angle of just two degrees, the XPD decreases from infinity to less than 30 dB. If both ends of the link sway in the opposite direction, the XPD equals to 23 dB which can already have an effect on some dual-polarized systems [87], [99].

Pitch γ_z – the rotation around the lateral z axis – is studied next. The situation is the same as before: two nodes are mounted on top poles of height h placed r apart on an even surface. This situation is illustrated in Figure 44. It is assumed that $r \gg h$, and that the antennas are very directive meaning the beam is at most couple of degrees in width. The radius of the illumination disc R is thus assumed to be smaller than the pole height h, i.e., $R < h$.

Pitching causes the AP move in the xy plane as shown in the Figure 44. As long as the target remains illuminated by the other endpoint, the link remains fully functional. The main beams are shown in red and blue. Neglecting side lobes, a suitable approximation for the antenna beam is a cone. Due to narrow beamwidth, however, the beam may considered a cylinder near the target.

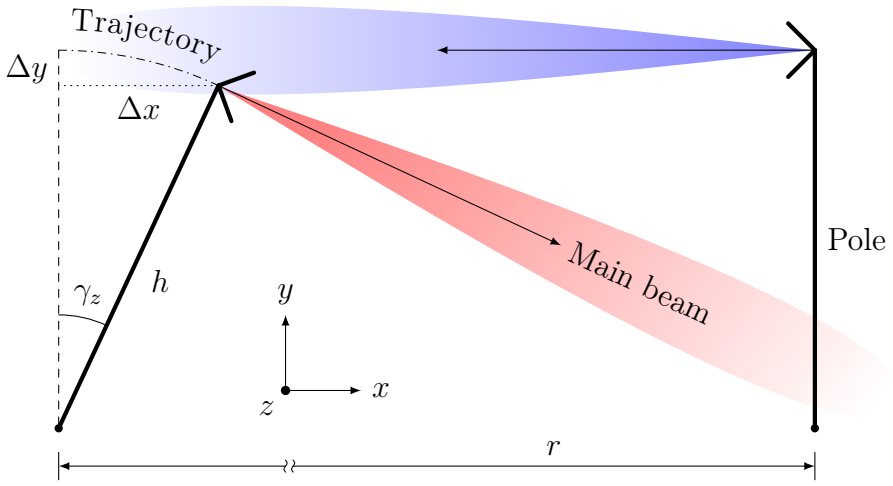


Figure 44: Misalignment caused by pitch γ_z.

Pitching causes the AP to move on a circular trajectory drawn using a dashdotted line in Figure 44. Two offset components may be identified. Even though greater of the two, longitudinal offset Δx is negligible compared to the link distance r since $\Delta x \leq h \ll r$. Moreover, longitudinal offset cannot cause misalignment when the endpoints are installed at the same height.

Only the vertical offset Δy is thus of interest, found using basic trigonometry

$$\Delta y(\gamma_z) = (\cos \gamma_z - 1)h. \tag{30}$$

Similarly to the case with roll, the offset Δy is always downward. Large pitching angles can cause the device to leave the illuminated area. This happens when

$$\gamma_z > \left| \arccos\left(1 + \frac{\Delta y}{h}\right) \right|. \tag{31}$$

Assuming a cylindrical target-centered beam, the link remains functional when the offset is less than R given by Equation (23). Using $R = 87$ cm from the previous example yields $\gamma_{z\max} = 27.01\,^\circ$. Since $\Delta y \leq 0$ and $\Delta z = 0$, the allowed pitch may be maximized by directing the beam R below the target. Now the maximum offset and corresponding pitch equal to $2R$ and $49.36\,^\circ$, respectively. Previously, the roll limit was maximized by targeting the beam 48 mm below the target. In this setting, the pitching limit equals to $27.75\,^\circ$.

Similarly to roll, it is highly unlikely pitching would cause a vertical offset with any noticeable effect on the link. Now the sway limits are even greater, and placement optimizations are not needed even though they increase the pitching range noticeably. Again, these ranges are rather optimistic since they assume precise alignment. However, pitching is not as susceptible to imperfect placement as roll since there is no lateral movement.

There are again two misalignment mechanisms; the displacement discussed above and antenna tilting which is more significant of the two. This is simply because an

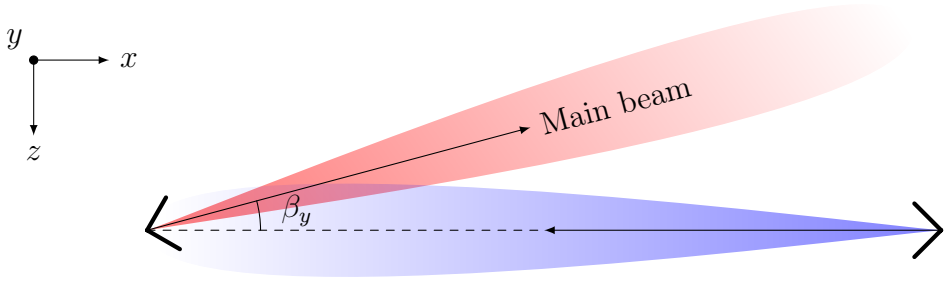


Figure 45: Misalignment caused by yaw β_y.

antenna tilt greater than half the *elevation* beamwidth causes the AP to miss its target. The example setup used symmetric antennas with one degree half-power beamwidth meaning the limits are half a degree in either direction. These limits are considerably stricter than those calculated before and easily exceeded in practice.

The tilting behavior is independent of both the pole height and the distance between the APs and does not affect alignment at the other end. The effect is confined within the affected access point, and can be mitigated internally using an orientation sensor (OS). This topic is discussed further in Section 6.3.

There are clear similarities between the two previous cases. In the presence of sway, i.e., roll, pitch or both, an AP mounted on top of a pole moves on a spherical $\alpha_x\gamma_z$ plane with a radius of h. The resulting displacement is solely because of the pole – more specifically the *leverage* measured from the pivoting point. Even without this leverage, roll induces polarization coupling while pitching steers the antenna vertically.

Lastly we consider yaw β_y – the rotation around the vertical y axis – using the same example with two APs mounted on top of two poles. This time, however, neither the pole length nor the link distance make any difference. In this setting, the pole twists around its center axis with negligible leverage. As a result, the AP stays exactly in its place as shown in Figure 45.

The effect is identical to the tilting phenomenon caused by pitch but in a different plane and without any movement. Similarly as before, the AP will no longer cover its target if it turns more than half its *azimuth* beamwidth. The other end remains unaffected, and the effect can be compensated using an OS as is discussed in Section 6.3.

6.2 Real-life characterization

In this section, misalignment is characterized by studying twist and sway statistics of prospective mounting structures. The presentation is based on extensive measurements reported by Kalimulin *et al.* [44], [50]. Conducted in 2014, these measurements took place in South-Western Russia and Southern Finland. Six different structures were investigated: a billboard and five poles or masts with different characteristics.

Table 9: Peak-to-peak sway and twist with 99.999 % probability [50].

Mounting structure	**Sway**	**Twist**
Billboard	8.0 °	1.6 °
Thin metal pole	7.6 °	1.0 °
Thick metal-concrete pole	2.5 °	0.6 °
Traffic lights	2.5 °	1.5 °
Telecommunications mast	2.9 °	1.2 °
Lamp post	3.6 °	1.2 °

The measurements were carried out by mounting an orientation sensor as if it were an access point. In addition to the OSs, anemometers and thermometers were included. Only the angular data from an OS was used, meaning movement was not considered. Moreover, the angular measurements were limited to pitch and yaw – roll was not measured. The same results may, however, be used for both roll and pitch since both are caused by sway.

The measurement results are summarized in Table 9 for each of the six different structures. The values are given as the maximum peak-to-peak range in degrees. They correspond to a probability level of 99.999 % meaning the deflection exceeds this range 10^{-5} of the time. It should be noted that the given absolute ranges are not necessarily symmetric with respect to the origin. The twist is approximately equally distributed whereas sway can be offset up to $2:1$ due to directional wind.

The billboard experienced most twist and sway due to its large canvas acting similarly to a sail. The board itself is also quite light and does not require rigid supports. As a result, it exhibits sway and twist amplitudes of 5.3 ° and 0.8 °, respectively. The thin metal pole behaves similarly whereas more rigid structures show smaller ranges and amplitudes.

In addition to angular statistics, also the timing behavior is of interest. For this, the sway angle was measured every 10 ms. Deflection times corresponding to a $\pm 0.2\,^\circ$ change in sway were measured for different structures. Depending on the structure, the $\pm 0.2\,^\circ$ deflection times were greater than 13 … 67 ms 99.999 % of the time.

All structures exhibit both fast and slow swaying, corresponding to deflection times ranging from tens of milliseconds to several minutes. Furthermore, both periodic phenomena and one-time events were recorded. The most common cause of both sway and twist is wind which has a clear effect regardless of the structure. The effect is dependent on both the mechanical structure and wind characteristics such as speed, direction and gustiness. The effect of thermal expansion and contraction was also evident. In addition to weather effects, some of the sway and twist were determined to have emerged from surrounding human activity, such as heavy traffic.

Table 10 compares different deflection mechanisms in the example used throughout the previous subsection. The maximum sway amplitudes given above are clearly

Table 10: Comparison of different deflection mechanisms.

Effect	**Movement**	**Rotation**	**Model**	**Measurement**
Lateral sway	Vert. + Horiz.	Pol. tilt	6.3 °	5.3 °
Longitudinal sway	Vert. only	Vert.	0.5 °	5.3 °
Twist	No	Horiz.	0.5 °	0.8 °

smaller than the displacement limits found for roll α_x. However, both the pitch and yaw amplitudes are large enough to have an effect on the quality of a fixed link. In the report [50], the simulated gain degradation exceeded 20 dB at reasonable probability levels. Sway and twist can thus reduce link quality considerably and even cause outages [32], [44], [50], [52], [87], [99].

Even though the measurement campaign undertaken by Kalimulin *et al.* [50] has been extensive, the values presented above do not cover all situations. After all, only six different mounting structures were investigated, and only in two different locations that are somewhat similar to one another in terms of climate. Additionally, channel effects in different deployment scenarios need to be considered. It is nevertheless certain that active misalignment compensation is an essential feature in high-density mmW mesh networks [30], [32], [44], [50], [51].

6.3 Maintaining precise alignment

Different means of improving the quality of existing links are considered next. Having completed the network expansion routine discussed in Section 5, the nodes have obtained initial alignment. Both nodes are properly synchronized and otherwise completely configured. The nodes may exchange data – at least until their antenna beams are no longer aligned due to twist, sway or both.

Within this thesis, the term *misalignment* is used to refer to small deflections due to weather effects and other phenomena. Broadly speaking, there are two different types or causes of misalignment: movement and rotation. Even though the two situations are similar, there are some differences. Regardless the type or cause, the compensation concerns only active links. Typically concerned with a single link at a time, the process may be optimized for each pair of nodes.

The following discussion is limited to such a situation, and further considers only a single data stream. The situation is somewhat different, i.e., more constricted, if the nodes use multiple streams to communicate one another. Moreover, there are some additional restrictions if one or both of the nodes communicate with multiple nodes simultaneously. For instance, to avoid TX-RX coupling in TDD systems, there can be limitations regarding simultaneous transmit and receive using different beams [83], [84], [88], [89], [99], [167], [168].

Before discussing *how* to approach the misalignment problem in different situations, we first address another important question: *when* to act. There are two

possible approaches; the active solution and passive operation. The former refers to maintaining the link fully operational in a proactive manner by performing the realignment procedure at regular intervals on a fully operational link. The problem with this approach is to find a suitable training interval. On one hand, increasing the alignment period increases both the chances and duration of a disconnect. On the other hand, too short a cycle results in high overheads. It must match the *beam* or *channel coherence time* [52], [57], [89] – a quantity related to deflection time discussed in the previous subsection.

In the passive approach, the need for realignment is determined on-demand, for instance upon detecting movement. Another possibility could involve tracking the signal quality over time and determine if realignment is in order. This approach can adapt to the situation at hand, and the actual channel coherence time instead of a crude estimate. While this can significantly reduce overheads, great care must be taken not to let the link disconnect. Not only does a disconnect create large delays, it would also have a serious impact on the entire network as it would trigger route recalculation [162]–[164]. In the extreme case, the node might connect to the network only through the link in question, meaning the AP would be dropped from the network resulting in even longer recovery delays.

The on-demand approach can provide better performance than the periodic approach. The question that remains, however, is how to schedule misalignment compensation dynamically on the fly. One possible solution would be mimic the periodic approach, and assign realignment slots periodically. These slots would be numerous but *conditional*, to be used only if necessary. Otherwise the link would function as normal. Such an arrangement would be regular and thus predictable but still capable of adapting to the current conditions.

The realignment slots would be used in two situations. First, either endpoint may request, or rather *order*, realignment an upcoming slot. To avoid excess delay, this request should always be followed without question. Secondly, a realignment is automatically scheduled in the next available slot if disruption in traffic is detected from missing packets. However, since only the receiver can detect missing packets, the realignment process should continue until alignment has been restored.

The most straightforward approach to reacquiring alignment would be to use exhaustive search where all possible beam pairs are tested. While the method is simple and guarantees optimal link quality, it is also very inefficient and likely to cause unacceptable delays. Furthermore, the effective search area cannot be reduced by increasing the beamwidth to speed up the process. This is simply because we are interested in fine alignment.

There are, however, two complementary approaches to reducing the training effort. First, the need for training can be greatly reduced or even completely removed using an inertial sensor and/or prediction schemes. Second, all pairs need not be tested if we accept any beam pair above some dynamic quality threshold rather than

trying to find the optimum pair. The *threshold-driven* approach thus considers not only signal quality but also the training delay when optimizing the link.

Next we discuss how to apply these mechanics to minimize disruption in different misalignment events, starting with movement. As already stated in Section 6.1, all movement is relative, and is consequently seen identically at both ends of a direct line-of-sight link. Such statements cannot be made for NLOS paths, as reflections mirror the movement [141]. In any case, both ends need to be realigned if one of the endpoints moves out of the illumination area.

In principle, this movement can be measured using three accelerometers and further compensated if the setting is known. In practice, however, the accuracy of these measurements commonly leaves plenty of room for improvement [187], [188]. Even small measurement errors in acceleration $\ddot{e}$ can cause significant offset in position e over time. The two are related quadratically, i.e., $e \propto \ddot{e}$, since position s is obtained as a double integral of acceleration $\ddot{s}$ as follows

$$s(t) = s(0) + \dot{s}(0)t + \int_0^t \int_0^\nu \ddot{s}(\upsilon)\, \mathrm{d}\upsilon\, \mathrm{d}\nu, \tag{32}$$

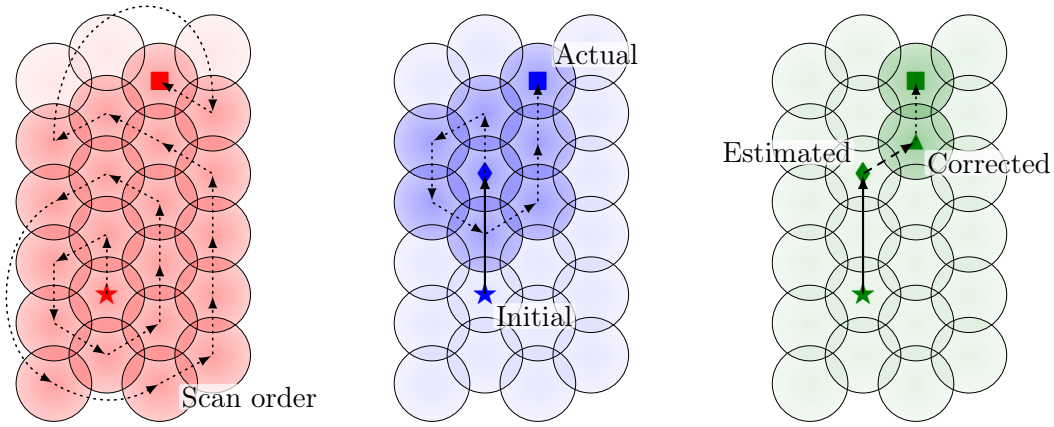
where $\dot{s}$ and t are velocity and time, respectively [199]. The previous holds also for angular quantities.

Even if accurate measurements were possible, they might not provide useful information since the hop length and the propagation path are not always known. Moreover, only the moving end would have its beams realigned while the opposite endpoint would need to be shared this information. This may prove difficult at large offsets using the mmW link as only the TX beam can be realigned before exchanging information, and the RX beam remains off-target.

The challenge with reducing the probing effort thus concerns efficient signaling. There are two cases worth noting, first of which is the situation described here where the millimeter-wave link is compromised. The mmW link should be kept operational by reacting to events without any excess delay. However, if the link is compromised as a result of sudden, large movement, there are a number of options to consider before probing.

If supported by the hardware, the system could increase the beamwidth and use a more robust modulation scheme when sharing the information. Depending on the gain roll-off of the misaligned beam, it might also be possible to simply switch the modulation scheme. The misalignment information could also be shared via a separate control channel already proposed in Section 4. The separate control channel could also provide useful in threshold-driven programming where the search is discontinued as soon as a suitable beam is found. The benefit of the entire approach is thus dependent on timely preemption of the search process.

To compare the *one-ended* probing effort of different approaches, consider the example shown in Figure 46 where an AP moves some beamwidths upwards and slightly to the right. In the figure, the antenna is assumed to have predetermined



(a) Blind: 21 (b) Sensor-assisted: 8 (c) Sensors & AI: 2

Figure 46: Different probing approaches with preemption.

4×6 beams laid out in a honeycomb-like pattern. The initial beam is marked with a star (★) whereas the target is denoted using a square (■).

The left pattern in Figure 46 corresponds to the *blind* approach where no sensor data is employed. The beams are scanned in an outward counterclockwise spiral pattern until current beam fulfills certain quality criteria. With the blind preemptive approach, 21 beam positions are tested. Unlike traditional exhaustive search, all 24 beams need not be tested since a suitable beam was found before the last beam. The probing order and tested beams are shown using dotted arrows and a darker background color, respectively.

Blind probing is naturally something to be avoided. While accurate movement estimates may not be available, one can determine the coarse direction of movement and possibly even its approximate magnitude [187], [188]. While this does not provide a complete solution even for the moving node, it provides a good starting for reducing the probing effort. Even though the information is not entirely reliable, the probing effort can be greatly reduced. To account for estimation errors, the search area should include large enough perimeter around the estimated direction.

This concept is elaborated in the middle of Figure 46. The situation is identical to the one used to demonstrate blind probing except for the sensor data. A compensation estimate based on accelerometer readings is now available. This estimate is shown using a solid arrow. The search begins at the estimated target marked with a diamond (♦). Using the same spiral probing pattern as before, the target beam is found on the eight try – a considerable improvement.

In an elegant system, accelerometer readings and resulting realignment needs would be recorded to improve the accuracy of future forecasts. This idea is shown on the right in Figure 46 where the sensor estimate is corrected using knowledge from previous runs. This correction factor is shown using a dashed arrow. The probing process is started from the beam marked with a triangle (▲), and since this is just below the target, only two training messages are needed.

Such a system would need to exhibit memorization and self-learning features: *pattern recognition* and limited artificial intelligence [200], [201]. While helpful sensor readings are not exactly necessary since the system can function based on alignment history. However, the more data there is available, the more reliable the forecasts. In addition to sensor readings and past alignment actions, the AP could also query how its movement is seen at the other end. While the active link is of prime concern, information from other nodes in different directions might prove useful in some situations.

The more accurate the corrected direction, the faster the whole alignment process. It should be noted, however, that Figure 46 is somewhat biased in favor of the advanced methods. To highlight the difference between the methods, the compensation need is rather large. In practice, the movement is typically small enough for one of the adjacent beams to be the target. However, large offsets are possible when switching between different links in a TDMA system. To avoid large switching delays, it is imperative to realign the link promptly. During a stable phase, however, benefits of these methods can be rather limited – especially if timely feedback is not available. Moreover, the effect may also be adverse if the estimate is clearly off-target.

Next we consider rotation which differs from movement in how the effect is seen at the other end. There are two types of rotational impairment: beam misalignment and rotated polarization. The former is caused by both pitch γ_z and yaw β_y while the latter is caused by roll α_x. As already pointed out in Section 6.1, rotation misaligns only the rotating end while the other end remains unaffected. The alignment may thus be completely corrected in an one-ended fashion without involving the other end at all.

Even though rotation causes misalignment more easily than movement, it is also more easily corrected due to its one-ended nature. Consequently, rotation information need not be shared unless it can be used to estimate movement. Other than that, the situation is similar and all previously discussed ideas apply. The need for compensation may be determined using orientation sensors, through probing, or by combining the approaches.

Using an OS can greatly reduce the alignment overhead but the estimation accuracy remains a question: how to minimize errors [187], [188]. Using relative measurements, the search area can nevertheless be reduced. The probing effort may further be reduced using adaptive compensation based on sensor readings and machine learning.

As already discussed in Section 6.1, rotation causes both polarization mismatch loss and coupling. The effect of the former was found negligible, and requires no actions as dual-polarized schemes complicate the hardware considerably. In CCDP systems, however, polarization coupling can be detrimental to overall performance [87], [99].

Knowing the roll angle α_x, a dual-polarized system could realign the polarization at both ends separately. If CCDP hardware is available and both α_x angles are

known at the TX, the system could pre-tilt the polarization for easier recovery. In practice, however, the benefits of such an arrangement would be limited since roll does not account for channel effects, such as multipath propagation and weather effects discussed in Section 3. To combat these effects, the system would need to incorporate cross-polarized interference canceler (XPIC) functionality [87], [99]. Regardless, polarization feedback might prove useful for optimizing link performance.

In the former, different types of misalignment were investigated separately. In practice, such categorization is not as straightforward since everything happens simultaneously. This is must be taken into account when exchanging position information between the endpoints. An AP reporting movement should not mistake its own rotation for movement at either end and vice versa.

To extend on the topic, it is further impossible to separate misalignment compensation and new link establishment from each other. The discussion on network expansion in Section 5 argued for maintaining a record of paired nodes. The node directions are recorded as they are seen from the mapping AP, and the results are thus dependent on both the position and orientation of the devices during the mapping. Moreover, the inter-node directions are not accurate since different nodes are mapped at different times and possibly under different effects. Given a dynamic situation, the endpoint may not be found where it was first recorded, and the situation is similar to what is shown in Figure 46.

As discussed in the previous subsection, the link is most susceptible to rotational offset. Most of the mapping error is due to orientation of the scanning AP. It is vitally important to remove the effect of any present misalignment components, and to record the direction where the target would be found if the AP was at rest. The node directions are thus *normalized* with respect to resting position. Before using the direction information, the mapping should in turn be denormalized, correcting for current misalignment.

This solves the aforementioned problem but only if the path is always found in same approximate direction. That might not be the case if the AP is subject to considerable movement. In essence, this minimizes the effect of misalignment local to the node. Moreover, if an AP can characterize the *movement* of other nodes, it can predict where the nodes can be found at a later point in time. This could be a great benefit in mesh networks where the nodes alternate the active link between multiple endpoints. Done right, this whole adaptation process could reduce initial locking time without wasting scarce resources – especially when exposed to periodic misalignment.

All in all, accurate orientation and position information, accrued from sensors and through self-learning, is advantageous for efficient operation. It should be noted, however, that low-end APs might not incorporate these features. In this case, the process relies solely on probing. However, it might still prove useful to exchange information, especially if it does not disturb actual data traffic. It is also possible

to partly outsource the decision making to the other endpoint if it supports the high-end features.

This might be the case with access links being rather dissimilar to backhaul links. First, the access network forms a star with the AP in the middle [83], [84], [100]. Secondly, the user equipment is limited not only in terms of hardware resources, e.g., processing power and sensors, but is also battery-powered [10], [16]. As a result, the AP would be well-suited for managing all access links, optimizing the system as a whole [115], [117], [118].

Furthermore, the misalignment characteristics of an UE are also rather different from of an AP [10], [16], [72]. While APs are stationary, users are free to use their device in any way, in any position and in any orientation. Even though people tend to use their phones in a certain way, user behavior is individual and effectively random. As such, it cannot be characterized in the general case – especially over long periods of time.

Consequently, access links have less stringent availability goals, and use antennas with larger beamwidth [10], [16], [72]. A wider beam can also be used in backhaul links that would be difficult to keep operational by other means. Narrow high-gain beams are naturally preferred since they provide better throughput with limited exposure. Throughput is decreased if the link has to compensate for smaller antenna gain by using *adaptive coding and modulation* (ACM) techniques [83]–[85], [87], [99], [118]. Commonly employed in modern telecommunication systems, the SNR requirements may be relaxed by reducing either the modulation order, the coding rate or both.

Wider beamwidth is justified only if the cost of maintaining the link using narrow beams is greater than that of a less efficient modulation scheme. While wider beamwidth certainly reduces gain and thus throughput, it can also improve availability. An example of a situation where a wider beam can be the best option involves long-term vibration with a large amplitude and/or short period, causing repeated runs of the alignment algorithm. This in turn causes large alignment overhead reducing throughput and increasing latency. Moreover, this can trigger a routing update if the service quality is compromised due to a congested or unreliable link [162]–[164].

7 Existing algorithms

In this section, a range of different beam steering algorithms are showcased. The presentation is general and sticks to a high abstraction level while concentrating on concepts and ideas rather than on actual program code. Moreover, the algorithms presented here do not necessarily correspond to the target system described in Section 4, and may be designed with entirely another application in mind. As a result, the algorithms do not necessarily take into account anything presented in Sections 5 and 6.

The presentation provides insight into three broad beam steering approaches and the current situation in terms of existing algorithms. First, Section 7.1 considers beamforming in the digital domain. This is followed by discussion on pure-analog solutions in Section 7.2. Finally, Section 7.3 investigates how these approaches may be combined in hybrid beamformers.

7.1 Digital solutions

In digital beamforming, multiple transceivers are used to digitize signals from different antenna elements separately [55], [57]–[60], [67]. As stated in Section 2.2, all functionality is implemented by means of digital signal processing. No analog precoding is employed, and the pattern is controlled by setting relative amplitudes and delays appropriately. As is typical in digital beamforming context, we consider only a single-ended system. On top of this, the focus is on reception since transmission is trivial after the target has been found.

Adaptive beamforming algorithms may be divided into two groups: those that use a reference signal and those that do not [57]. Algorithms in the latter category involve two components: direction of arrival (DoA) estimation and pattern generation. Both of these can, in turn, be achieved in a number of ways. Regardless the approach, digital beamforming can be considered as a type of channel estimation best known from spatial multiplexing [89], [90].

There are two main approaches to DoA estimation: conventional methods and subspace-based algorithms [57]. The *conventional methods* are based on a power spectrum obtained using the beamforming concept discussed in Section 2. The environment is scanned in all possible directions, and sources correspond to maxima on the power spectrum. There are two well-known implementations; the classical *delay-and-sum method* and *Capon's minimum variance technique*.

Also referred to as the *Fourier* or *Bartlett method*, conventional beamformers work identically to a codebook-based analog beamformer [57]. Determined by the used codebook, the angular resolution of this method is usually quite poor. If there is no correlated interference, the angular separation may be improved using *minimum variance distortionless look* commonly known as Capon's technique [57]. In this approach, peaks on the power spectrum are sharpened by minimizing the output power while keeping peak power constant.

Subspace-sampling methods take advantage of both the structure and statistics of the received signal [57]. Consequently, they achieve considerable performance improvements over all conventional methods. The approach is based on the notion that signal and noise subspaces are orthogonal to one another. The two most popular DoA estimation algorithms, namely MUSIC and ESPRIT, are based on the subspace approach [57]. These are discussed next.

First proposed in 1979, MUSIC stands for *MUltiple SIgnal Classification* [57]. In the algorithm, the signal and noise subspaces are distinguished using eigendecomposition on the received signal covariance matrix. The MUSIC may be applied to any array geometry, and is capable of determining multiple parameters of the radio environment, such as the number of sources, their DoAs, relative strengths and cross-correlation, and noise powers.

On the other hand, there are some drawbacks [57]. First, the array configuration and response need to be accurately known, and storing this information requires considerable storage space. Second, the algorithm is computationally demanding, and the estimation process is exhaustive, meaning only one parameter is resolved at a time. Third, the algorithm does not perform well in low-SNR situations or when the sources are closely spaced. Lastly, the algorithm is unfit for multipath environments since it fails when the impinging signals are correlated.

Some of these drawbacks were later addressed in subsequent implementations such as Root-MUSIC, Cyclic MUSIC or Fast Subspace Decomposition [57]. Still, better performance is achieved with the *Estimation of Signal Parameters via Rotational Invariant Techniques* (ESPRIT) algorithm developed in 1989 [57]. Compared to MUSIC, ESPRIT is not as computationally intensive nor does it require as much storage space. The algorithm does not involve exhaustive searching, and the array need not be accurately calibrated. All this is possible when the array has a *displacement invariance* property meaning the elements in the array form matched pairs with equal displacement between the paired elements. Not only does this facilitate computation considerably, but the condition is met with relative ease in practical arrays.

After obtaining the DoA estimates, the pattern may be generated. The simplest approach is simply to steer the beam maximum toward the target similarly to a phased antenna system [57]. Interferers are ignored meaning pattern minima are not placed in these directions. This approach is commonly referred to as the *classical beamformer* and mirrors the delay-and-sum DoA estimation method. It can be extended to maximize SNR if the noise covariance matrix is known [57]. Known as a *maximum SNR beamformer*, the technique corresponds to the classical beamformer if the noise is assumed to be spatially white.

Interferers may be suppressed using an approach like the *multiple sidelobe canceler* (MSC) method where the antenna weights are obtained through a two-step process [57]. First, the system generates patterns that would separately maximize each source using conventional beamforming. The final weights are then found as a suitable

linear combination of these source-specific weights that minimize off-target gain. A suitable pattern maximizes the achieved signal-to-interference and noise-ratio (SINR). Because of this, the method is also referred to as the *maximum SINR beamformer*.

Next, algorithms that have prior knowledge of the desired signal are discussed. Naturally the actual data is not known beforehand but some of the desired signal characteristics, such as its power envelope, may very well be. It is also possible to use predetermined training signals for this purpose. The beamforming vectors are found by recovering this property or the pilot from the received signal [57]. The recovery procedure involves optimizing some performance measure, and such algorithms are often referred to as *statistically optimal beamformers*.

Some of the typical optimization approaches include minimizing mean square error (MMSE), maximizing SNR (MSNR), and minimizing noise variance (MNV) [57]. These are only quality metrics, and the algorithm implementations may vary. A simple approach to the MMSE problem would be direct matrix inversion which does not enforce any constraints on the beamforming vector. Usually there are, however, certain practical constraints, and one might, for instance, use the *linearly constrained minimum variance* (LCMV) approach [57]. This method is somewhat related to aforementioned Capon's technique, and may be adapted to maximize signal likelihood.

Employed in numerous telecommunication systems, the most widespread adaptive beamforming algorithm is the *least mean squares* (LMS) algorithm [57]. This is because of its good error tolerance and low computational complexity. The algorithm uses *negative steepest descent method* to estimate the change gradient, and to update its beamforming vectors accordingly. While the method is guaranteed to converge, the convergence may take arbitrarily long in certain conditions.

A number of LMS variants have been proposed to address this issue [57]. One of such algorithms is the *recursive least squares* (RLS) algorithm which utilizes input history. While the RLS indeed diverges much faster than the LMS method, there are some drawbacks. Most importantly the RLS algorithm is significantly more computationally intensive than LMS. In addition to LMS, RLS or other LMS variants, there are also a number of other approaches, such as *constant modulus algorithms* (CMA), *affine-projection methods* and the *quasi-Newton method* [57]. Some of these mechanisms, such as the CMA, can also be combined with the subspace-based DoA estimation methods.

Focusing on analog methods and a specific application, the discussion on digital beamforming has been necessarily short and incomplete by far. Digital beamforming has been extensively researched in the past few decades, and the findings have been published in numerous publications. The interested reader is referred to textbook material on the topic; see Balanis's books [54], [57] for a comprehensive introduction and Liu's book [59] for more advanced discussion.

7.2 Single-transceiver algorithms

Pure-analog methods and systems comprising a single transceiver, such as analog beamformers, are considered next. In analog beamforming, signals from different antenna elements are both weighted and combined in the analog domain before digitization [52], [54], [55], [58], [63], [67], [81]. Since an overwhelming majority of practical analog beamforming systems employ predetermined weight vectors, the following discussion is limited to such systems.

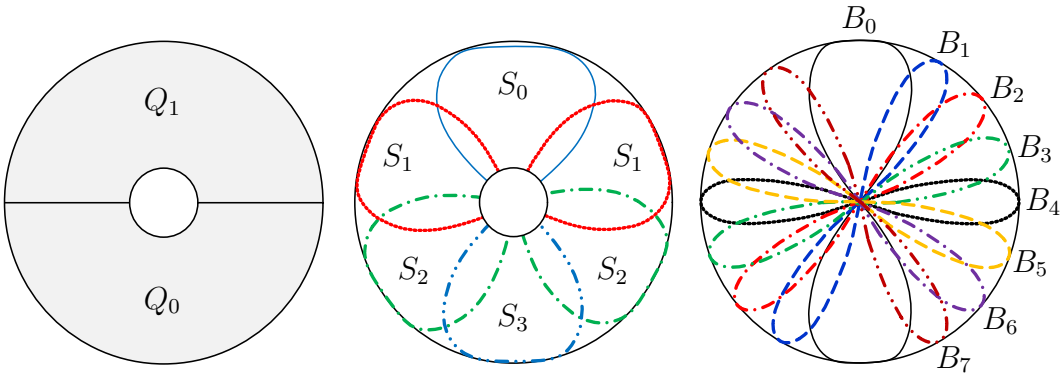
As is typical to most practical ABFs, the investigated systems do not involve gain control for complex pattern generation [55]–[58], [60]. Still, many implementations involve beams of different width and gain using On/Off switching. Among the showcased systems, only a single example employs variable gain control, and only for windowing purposes. On the other hand, there are also a number of pure beam switched approaches. Regardless the mechanism, interference suppression is not considered.

Since there is but a single transceiver, complex channel estimation methods employed in digital beamforming systems cannot be used [55], [57], [58], [60]. However, classical beam searching is possible. The exhaustive search used in Sections 5 and 6 forms the basis of all single-transceiver algorithms. While considerably less evident than in digital beamforming methods, beam searching is also a type of channel estimation. It is simply enough to know the directions in which the channel is most favorable for transmission to a certain endpoint.

As argued above, all single-receiver methods involve a type of beam search. There are, however, two different approaches to conducting such a search. The straightforward method is to test all possible directions in a hit-or-miss fashion as is done in exhaustive searching. Since this is quite time-consuming, it is common to limit the effective search area by using wider beams. The entire search area is nevertheless covered. The other technique uses previous measurements to choose tested directions in such a way that the entire search area need not be covered.

Conventional methods are discussed first. The exhaustive search is simple, robust and guaranteed to work even in challenging conditions with multipath propagation [57]. Included in a number of short-range mmW telecommunication standards, it is also mature and proven in practice. In the following, overview the beam steering aspects of IEEE 802.11ad [40], IEEE 802.15.3c [41], ECMA-387 [42] and WirelessHD [43] standards are presented. This is followed by descriptions of a number of other algorithms intended for mmW cellular communications.

The IEEE 802.15.3c standard [41] defines an alternate 60 GHz physical layer to the high data rate wireless personal area network (WPAN). Optional beamforming functionality is defined to support multiple different antenna configurations: beamforming arrays, beam switched antennas, and even fixed antennas. While optional, beamforming improves both range and throughput. Two protocols are specified: on-demand and proactive beamforming. They are almost identical but target slightly different situations.



(a) Quasi-omni (b) Sector (c) Fine

Figure 47: IEEE 802.15.3c codebook proposed by Wang [69].

Originally presented by Wang *et al.* [69], the beamforming functionality targets simple and low-cost hardware [75]. As a result, the array is either uniform linear or planar array (ULA or UPA, respectively) with half-wave spacing between the elements. The codebook is typically constructed using 2-bit phase shifters with 0 °, 90 °, 180 °, and 270 ° as the shifting options. An example of an IEEE 802.15.3c codebook with a four-element antenna was given in Figure 4 during the discussion on codebooks in Section 2.1.

In addition to phase shifting, On/Off switching is employed to generate beams of different resolution. As shown in Figure 47, there are three types of patterns, called quasi-omni, sector and fine beam patterns. Quasi-omni patterns are few in number and have low gain but are in turn wide enough cover multiple sector-wide beams. A single sector-wide beam further covers a number of precise beams used for data transfer. The quasi-omni and sector-wide beams are only used to improve the setup time via hierarchical searching. Not specified in the standard, their characteristics are left to implementation. The standard codebook thus specifies only the high-resolutions beams.

The beam alignment procedure comprises three stages: device-to-device linking, sector-wide search and beam-level search which correspond to the three types of beams. All stages are similar in structure, and consist of training, feedback, mapping and acknowledgment subphases. In the training phase, the receiver decides the optimal TX/RX beams of current resolution by estimating SINR from received training sequences. The results of this exhaustive search are fed back to the transmitter during the feedback phase. In the mapping subphase, the endpoints exchange information regarding how the best beam maps to finer beams used in the subsequent stage. Finally, the two ends acknowledge before moving to the next stage.

In addition to aforementioned initial alignment, the IEEE 802.15.3c standard [41] defines optional tracking functionality to mitigate the effects of a time-varying channel. In tracking, the currently used beam and some of the neighboring beams are grouped as cluster of specified shape. Beams in the cluster are tested periodically, and a beam

is switched if the estimated SINR improves. If the link cannot be maintained using one of the beams in the cluster, the link times out and is dropped.

Being one of the first widespread standards to involve beamforming, the IEEE 802.15.3c has seen its share of controversy regarding its codebook design and resulting slow setup. For instance, Feng *et al.* [68] suggest using a two-tiered uniform circular array (UCA) with $8 + 8N$ ($N \in \mathbb{N}^+$) elements instead of a ULA. Used to create both the quasi-omni and the sector patterns, the 8-element inner-tier would be fixed across all implementations. Consisting of $8N$ elements, the outer tier is used to create equispaced beams of equal gain and with identical shape.

There are two quasi-omni patterns and eight sector-wide beams. The quasi-omni pattern has the shape obtained by superposing every other sector beam. Such an arrangement streamlines the search, and Feng *et al.* propose single-sided searching. Only the quasi-omni pattern is searched in an exhaustive manner, and subsequent stages simply align one end at a time. As a result, the search is no longer quadratic but linear. The UCA layout also improves both antenna gain and accuracy. Unfortunately, the method is not as robust and the UCA layout cannot easily be extended to another dimension.

The beamforming functionality of the ECMA-387 standard [42] is similar to that included in the WPAN standard but simpler. It implements both initial acquisition and tracking features using a pairwise search in two operation modes. In the closed-loop mode, the TX may request the RX to feed back the optimal transmit weights. The open-loop configuration is based on reciprocity, and the TX uses best receive vector also for transmit.

The antenna patterns are generated using either a Fourier or Hadamard codebook [42]. Also referred to as a DFT codebook, a Fourier codebook requires as many equispaced phase shifting options as there are beams [65]. Hadamard codebooks, on the other hand, are created using up to four phase shifting values similarly to the 60 GHz WPAN codebook. The Fourier codebook is more complex but performs better when there are large number of beams or antenna elements. In any case, the beams are all of similar resolution since only phase shifting is employed.

The WirelessHD 1.1 [43] specification is similar to the ECMA-387 standard but leaves more room for custom implementation. There are again two operation modes for initial beam search and tracking. Like before, there are two feedback modes, namely explicit and implicit feedback, which operate similarly to the closed and open-loop configurations of ECMA-387. While the standard is primarily intended for codebook-based ABFs, it also incorporates spatial multiplexing functionality using hybrid beamformers.

Similarly to IEEE 802.15.3c, also the IEEE 802.11ad [40] specification is a 60 GHz amendment to the IEEE 802.11 wireless local area network (WLAN) standard family. The specification naturally includes beamforming but only in terms of general framework and messaging. Antenna specifications and the actual beamforming

algorithm is left to implementation. In the following, the beam alignment procedure is outlined followed by an example implementation by Zhou *et al.* [65].

The IEEE 802.11ad [40] beamforming framework is similar to that defined in the IEEE 802.15.3c standard [41]. The framework is two-fold with separate procedures for beam acquisition and beam tracking. The initial alignment involves three stages, namely sector-level sweep, multiple sector identification and beam combining. The first two stages provide initial beams which are later refined in the last stage also referred to as the beam refinement procedure.

The process starts with the sector-level sweep where the initiator transmits while the responder listens. During the phase, the TX uses narrow beams while the RX employs a broad quasi-omni pattern. The TX tests as many precise beams as it wishes by sending training messages before both the roles and patterns are switched. The initiator receives using a wide beam while the original responder probes its high-gain beams.

After obtaining the initial transmit beams, there are two options for obtaining the initial receive beams. The APs may either use the TX beams based on reciprocity or perform a similar search to find RX beams. The optional multiple sector identification phase is there for this purpose. The phase is otherwise identical to the sector-level sweep except the roles are now reversed. The TX uses a quasi-omni pattern while the RX tries to receive using different high-gain beams.

Now both ends of the link have their initial beams. However, these beams may not be optimal for subsequent data transfer. This is because the quasi-omni pattern may have significant gain fluctuation across the main lobe. The final beams are determined in the beam combining phase where the most promising beams are tested in pairs. Within this set, the process is a two-way exhaustive search with fast-slow division discussed in Section 5.3.

As stated before, the standard also defines a tracking framework but the actual functionality is left to implementation. Only a training field is defined to carry beamforming related messaging as a part of the payload. This enables limited alignment functionality using the communication beams without interrupting the flow of data. While tracking clusters like those used in IEEE 802.15.3c [41] are not defined, the implementation may include similar functionality.

Zhou *et al.* [65] adopt the previously discussed DFT codebook in their phased-array implementation. While their implementation adheres to the IEEE 802.11ad standard, they propose a slightly different approach for beam refinement. Instead of choosing the refined beams directly from the initial set, their propose choosing three or five beams around the target direction from a separate, denser codebook. The authors propose a three-level scheme similar to IEEE 802.15.3c where the only the high resolution beams are used for data streaming. While this increases gain, it also requires more accurate phase-shifters.

It should be noted that all of the former solutions are designed to operate in an indoor environment, and over a very short range. The first three are defined up to a

few meters [41]–[43] while WLAN [40] is defined up to a couple of dozen meters. The range is not the only difference between aforementioned systems and mmW mobile networks; the setting, environment, specifications and goals are totally different [52], [120], [125]. Consequently, these solutions cannot directly be employed in wireless backhaul. Regardless, some of the principles may still be used. Next, a number of approaches intended for mmW mobile communications are showcased.

Hur *et al.* [63] propose using a multi-tiered approach similar to previously discussed WPAN and WLAN standards. For this purpose, they have created a hierarchical, tree-shaped codebook with beams of different widths at different levels. The proposed codebook is organized in K ($K \in \mathbb{N}^+$) levels, and a single coarse beam covers N ($N \in \mathbb{N}^+$) beams of finer resolution. The branching factor thus equals to N, meaning there are N^K beams of the highest precision.

Assuming identical and reciprocal devices at both ends, the (asymptotic) complexity of a direct exhaustive joint-search would be $\mathcal{O}(N^{2K})$. The corresponding complexity for a single-sided search and the proposed method are $\mathcal{O}(2N^K)$ and $\mathcal{O}(KN^2)$, respectively. As an example, consider a codebook with $K = 3$ and $N = 4$ shown in Figure 48. There are 64 narrow beams, and the previous complexities are 4096, 128 and 48, respectively.

Even though the algorithm itself is a typical exhaustive search of quadratic complexity at each level, the total complexity is dramatically decreased due to being a N-ary search [197]. The probing effort could be further decreased using a taller tree. For instance, by setting $N = 2$ and $K = 6$, the total complexity is 24. This would, however, be near-sighted since feedback overheads increase with the codebook depth K [63]. As a result, the required training time could increase.

Unlike before, the variable beamwidth is not implemented by controlling the number of active antenna elements. Instead, the wider beams are generated using multiple slightly offset beams using the squinting principle discussed in Section 2.1. While the variable beamwidth functionality is phase-controlled, Hur *et al.* employ

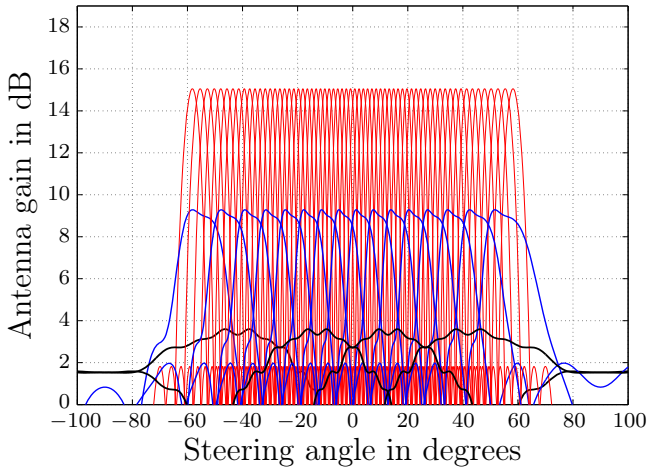


Figure 48: The 3×4 codebook proposed by Hur [63].

amplitude control. This is because they have decided to use Taylor windowing to focus the energy in the main beam, resulting in lower sidelobes.

In 2013 – two years after their initial publication – the same group published a series of extensions. They still employ a hierarchical codebook but use a different pairing method. First, Hur *et al.* [52] proposed an adaptive subsampling algorithm referred to as ping-pong sampling or the back-and-forth method. The alignment process includes multiple rounds with a serve and a return where the TX probes the channel with assistance from RX and vice versa. While this requires less probing effort, it also causes larger feedback overhead as the next step is determined from previous measurements.

Larew *et al.* [114] then adapted the proposed method to dual-polarized MIMO systems. The search space is halved by partitioning the codebook into sectors, and using two orthogonal polarizations in parallel in different areas. The simultaneously scanned segments are chosen to be clearly separated in distance. The ping-pong sampling method is still used for all rounds except for the first where conventional joint-search is employed. This is done to avoid errors due to cross-polar components.

Being a MIMO system with two independent streams, two transceiver chains are required at both ends of the link. As such, the system could also be considered a hybrid beamformer. Still, the two transceivers work independently using orthogonal polarizations, and cannot be used to contribute to a single stream. In essence, the system comprises two separate analog beamformers and can thus be effectively considered a single-transceiver system.

Lastly, Song *et al.* [115] adapted said method to multi-user scenarios. Naturally, MU-MIMO requires a full-fledged hybrid beamformer at the AP with as many transceiver chains as there are simultaneous users. However, in their proposal, Song *et al.* treat parallel analog beamformers as if they were separate. The algorithm works as the aforementioned hybrid method, and the users are divided into groups where inter-user distance is maximized to minimize cross-polarization errors. The beamforming vectors are chosen to maximize the total SINR between all users.

A couple of beam switching solutions are introduced next, starting with the work of Kalimulin *et al.* [184]. Their proposal is mostly built using the elements discussed above. For instance, the FDD algorithm implements both initial alignment and tracking functionality. Furthermore, the actual training function is a typical exhaustive joint-search with fast-slow division. Still, there are some features worth discussing.

During the initial alignment, the devices need not be synchronized with each other. Required time synchronization is first acquired during the search mode, and the realignment mode assumes proper synchronization. The realignment process is triggered using a RSSI threshold or a periodic timer. Implemented using Zadoff-Chu sequences and matched filters, the beam training will always test all possible beam conditions.

Last example of the hit-or-miss search algorithms showcased here was proposed by Veijalainen [202]. Yet again both acquisition and tracking are implemented using a two-sided exhaustive search in a fast-slow manner. Veijalainen is, however, the only one to employ an orientation sensor to counter device deflection. Moreover, the proposed system was actually implemented in practice whereas Hur *et al.* [52], [63], [114] and Kalimulin *et al.* [184] base their arguments solely on simulations.

Aimed for vibration compensation, the tracking functionality supports three operation modes. First, the orientation estimation method switches the beam when suggested by the orientation sensor if the measured RSSI is below a certain threshold. The second method does not employ the sensor but probes the neighboring beams periodically and switches to another beam if it yields better RSSI. The third method combines the approaches, probing the immediate beams only when triggered by vibration. To avoid excessive switching, the threshold must be exceeded during multiple consecutive measurements.

Next, we will consider the other school of single-receiver algorithms where the entire search area need not be covered [66], [72], [203]. The previous *traversal-querying* methods test all combinations blindly, only capable of recognizing a suitable solution when found. The key idea behind the more advanced methods is that all tests contribute to the final solution. The tested beams are determined on the fly based previous tests. Such a process is sometimes referred to as an *active search*.

While the probing effort is significantly decreased, the performance does not improve on a one-to-one basis [66], [72], [203]. This is because of two reasons. First, these algorithms require significant, albeit still realistic, amount of computation. But more importantly, there is a significant signaling overhead if the next step is determined on the fly based on the previous measurements. The conventional methods employ predetermined, systematic search patterns, and the signaling may be pipelined and sent in bursts. That is, multiple consecutive training messages are acknowledged using a single cumulative reply as discussed in Section 4.1.

It should also be noted that methods based on active search have gained researchers' attention only recently [66], [72], [203]. Consequently, they have not been proven in practice, and arguments for them are based on limited numerical simulations. This is the case even though the methods themselves are based on extensive numerical optimization theory. Moreover, these methods are targeted to be used for initial alignment while tracking is still best implemented using conventional means given the small search area.

Active search algorithms exhibit clear master-slave protocol. The algorithm is run only on one of the endpoints while the other end simply follows given instructions. While everything is orchestrated and controlled by the master, the slave may implement either the transmit or the receive-calculate-respond functionality. While the latter benefits from increased processing power, the first typically consumes more power. In any case, a dedicated control connection is needed in addition to the mmW link.

Considering the numerical optimization problem, the first task is to define the objective function [204]. This is also referred to as the reward or cost function depending if the optimal solution is the maximum or minimum, respectively. The algorithm may use any link quality metric that can be determined with relative ease. Measured from predetermined training sequences with favorable autocorrelation characteristics, signal-to-noise-ratio is a good candidate [66], [203].

All practical link quality metrics share four properties [66], [72], [203]. First, the current value of the objective function is rather expensive to obtain. Second, the reward function is *single-agent* in nature meaning a single measurement returns only one value of the objective function. Third, explicit derivatives and analytical gradient information are unavailable. Fourth, the objective function has multiple local optima as illustrated in Figure 49.

The numerical pattern search machine should thus employ only previous values of the reward function to approximate the gradient [66], [72], [203]. Given these constraints, Li *et al.* [66] suggest approximating the gradient using a *Rosenbrock search* in order to maximize received SNR. The routine involves two phases where either the probe or the pattern moves. The phases can be though of as peak finding and peak refinement, respectively.

The algorithm starts in the probe move phase after generating orthonormal direction vectors in all orthogonal directions. The total number of dimensions equals to the number of scanned planes, and is thus determined by the beam layout. If both ends of the link employ one or two-dimensional arrays, there are two or four orthogonal dimensions, respectively. One should also note that the beams can no longer be simply enumerated.

In the first phase, the probe is moved along these orthogonal dimensions, one at a time. The probe is moved forward a single step, and the value of reward function

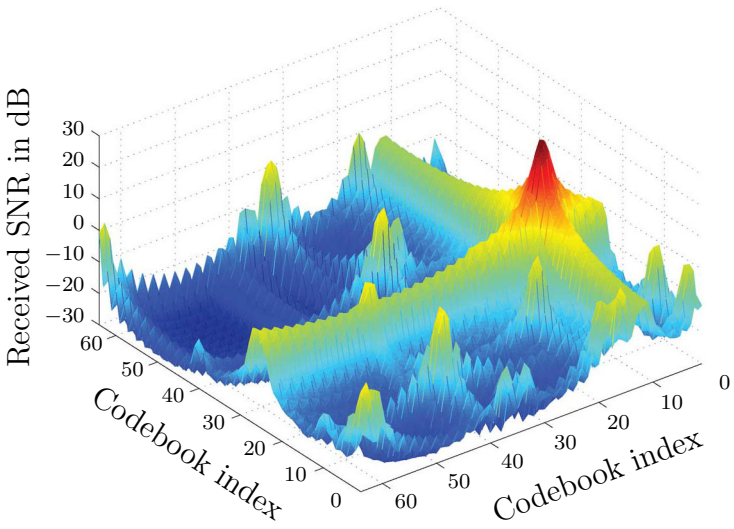


Figure 49: Received SNR with different ULA beam pairs [203].

is determined. If the value of the reward function has increased, the move was successful. This point becomes the new starting point and the corresponding SNR is recorded. The step distance is also increased using a multiplication factor. Upon a failure, the algorithm backtracks by starting again from the previous point but in the opposite direction with a smaller step size. This process is continued when the move fails in all directions.

The pattern move involves generating a new set of orthonormal direction vectors. For this, Li *et al.* use Gram-Schmidt orthogonalization. After obtaining a new set of direction vectors, the algorithm returns to the probe move phase. These two phases are alternated until the improvement during a single probe move phase is below a predetermined threshold.

Being a hill-climbing algorithm, the solution may only improve during the Rosenbrock search. Due to a non-smooth reward function with multiple local maxima, this greedy approach may return a locally optimal solution. As this is strongly dependent on the starting location and the initial direction vectors, Li *et al.* combine the Rosenbrock search with a initialization routine. This pre-search involves a conventional neighborhood search using limited number of beam pairs.

Even with the initialization, the algorithm occasionally fails to converge on the global optimum. In these cases the algorithm may be run again starting from a different point. In fact, the asymptotic probing complexity of the algorithm is dominated by the initialization while the Rosenbrock search is mostly unaffected by the search area. Considering an example with 64 beam options at both ends, the reported average probing effort was 38 compared to 4096 of the conventional exhaustive joint-search.

In 2014 – a year after announcing the original algorithm [66] – Li *et al.* [203] reported an improved solution combining *simulated annealing* with the Rosenbrock algorithm. Simulated annealing is a generalization of the Monte Carlo approach that falls within the biological algorithms category. In essence, the algorithm may escape local optima with certain gradually reducing probability.

This random behavior is implemented using two different random generators for the global and local scopes, respectively. While the local probability is derived from the global value, both decrease at geometric rate. The global probability is updated always whenever direction vectors are updated, and the updated value is used in the next probing phase as the local probability. In turn, the local value is updated always when probe move worsens the solution.

The current value of the objective function may thus be worse than the previous iteration. This is more probable in the beginning, and the algorithm reduces to a typical numerical search toward the end. While this makes it possible to escape local optima increasing the rate of success, search failures are still possible. On top of this, the algorithm requires more iterations to converge. Regardless, the algorithm requires less probing than those defined within the IEEE 802.15.3c WPAN [41] or IEEE

802.11ad WLAN [40] standards. However, the amount of feedback has multiplied many times over.

Finally, the algorithm developed within the MiWaveS project [72] is discussed. Intended for the system described in Section 4, the algorithm will be implemented using LabVIEW and run on the prototyping hardware. Among the outcomes of the MiWaveS project are an access link alignment algorithm, a backhaul steering solution, and an efficient codebook design using genetic algorithms. The access link alignment algorithm is discussed below.

The initial alignment functionality developed by Chiang *et al.* involves a numerical search similar to previously discussed Rosenbrock search. For all things considered, the setting is identical but the optimization method is different. The proposed deterministically sampling black box algorithm is a two-dimensional extension to the Brent's method and employs *line spline interpolation* techniques. The method is computationally more demanding but comes with reduced feedback overhead since the measurements are made in bursts.

Referred to as the *mode pursuing sampling method for a discrete variable space* (D-MPS), the process involves balancing discovered optima and possible new optima. This balancing act employs two circles with increasing and decreasing radii, respectively. Confined to the active area, line spline interpolation is used to model the received SNR without evaluating the actual objective function.

The algorithm starts by measuring a number of beam pairs within the active area. Referred to as expensive points, they are picked randomly according to a predefined pattern to allow equidistant exploration. The beam pair that maximizes SNR is chosen as the center of the two circles, finishing the initialization phase. Next, the D-MPS algorithm is ran until the optimum beam pair is found or the maximum number of iterations has been reached.

Within each circle, the same number of cheap points are picked randomly. Using line spline interpolation and previously measured points, an SNR surface is fitted and the objective function is estimated in the cheap points. After this, a small set of new expensive points are chosen and measured. From the set of all measured points, the best is chosen as the new center. If the new center was found among those just measured, the smaller and bigger circles are increased and reduced, respectively. If there have been no improvements during a predetermined number cycles, the radii are scaled the opposite way.

7.3 Hybrid beamforming

A hybrid beamformer comprises a small number of transceiver chains and element-wise weight elements [28], [37], [67], [76]–[78], [80], [92]. A hybrid beamformer attempts to compromise between flexible but costly digital beamforming and practical but restricted single-receiver solutions. The question is how to combine analog and digital signal processing to get best of both worlds.

There are two separate yet interlinked design tasks: the analog RF precoder and the digital BB combiner. At the receiver, signal recovery is difficult since the BB processor does not have access to raw element-wise signals like in conventional sub-6 GHz MIMO systems. Instead, only a small set of perturbed signals are available after weighting and combining in the analog domain.

This challenge is still relevant and evident in all HBF publications, most of which focus on the baseband functionality and the MIMO processor. In early HBF systems, the problem was addressed by allocating each transceiver chain with their own antenna [78]. Referred to as *antenna selection*, the method essentially converts the hybrid beamformer into a digital beamformer. Providing access to raw antenna signals, conventional MIMO designs may be used with minor modifications. On the other hand, the performance of such systems is poor due to small array gain.

It is also entirely possible to approach the issue from the opposite direction, resorting to pure-analog methods. A hybrid beamforming hardware may be used to implement an analog beamformer with minimal changes – the signals from different transceivers need only be added coherently. While straightforward, this does not take advantage of all available hardware. A better solution would implement multiple, separate ABFs in parallel similarly to what Song *et al.* [115] have done – see the previous subsection.

The analog beamformers can also be replaced by switched beam antennas that support multiple independent beams [205]. The resulting multi-transceiver switched beam system could be considered an extension to the aforementioned antenna selection method. The extension would, however, require a training scheme of some sort to first align the beams before the streams can be combined in the MIMO processor.

El Ayach *et al.* [205] have taken such an approach in their multimode precoder. Their hybrid beamformer comprises several disjoint subarrays with element-wise phase shifters. All precoding is carried out in the analog domain, and the digital BB processor is only used for stream-to-subarray mapping. If SNR is high enough, the system may allocate a subarray with a new, independent stream. Otherwise the system uses the subarray to improve the SNR of an existing stream. The process is iterative, optimizing one chain at a time through constrained quadratic maximization.

While most HBF publications focus on baseband processing and approach the topic from a digital perspective, most of the solutions are quite similar to one another. There are, however, a number of approaches that take a different stand. Such methods are discussed before showcasing a number of nearly identical solutions based on matching pursuit. The section finishes with a real-life example.

Huang *et al.* [77] investigate HBFs implemented using both interleaved and side-by-side subarrays. Angle of arrival (AoA) estimation is based on phase difference between the subarrays via conjugate multiplication and cross-correlation. The analog precoders are tuned *recursively* based on feedback generated as a result of this multiplication with increasing precision. Using FFT and its inverse (IFFT), all signal

processing is carried out in the frequency domain. Huang *et al.* argue for frequency domain processing because of its easier implementation.

Two algorithms are proposed, namely *differential beam tracking* and *differential beam search*. Neither of the algorithms requires knowledge of the reference signal, and are immune to Doppler shift due to a moving target. Computationally less intensive of the two, the tracking algorithm is primarily intended for interleaved arrays. The search algorithm is a side-by-side extension of the tracking functionality that scans the whole beamspace. The authors suggest running the search algorithm only once and running the tracking algorithm periodically thereafter. However, also the tracking algorithm may be used for initial acquisition.

Even though the simulation results show fast convergence, the approach is limited to narrowband signals. Furthermore, it is susceptible to mutual coupling between the antenna elements. A year later, the same group reported an improvement [92] to their original concept. The approach is otherwise identical but the phase difference is now measured at a certain frequency offset. For more accurate results, multiple measurements are made at different frequencies using large offsets.

Another interesting HBF solution has been suggested by Kim *et al.* [28] for MU-MIMO. What makes the approach interesting is that both the analog RF precoder and the digital BB combiner weights are taken from a codebook. Using a reference signal, the BS periodically measures the quality of relevant beams. The active beams are then chosen from the analog codebook using a *proportionally fair* scheme to maximize channel capacity. The digital combiner weight matrix is then determined using a zero-forcing algorithm to minimize co-channel interference based on feedback from different user terminals.

Next, the mainstream approach to hybrid beamforming is discussed. Based on manipulation of the MIMO channel, all these methods are not only similar to each other but also mathematically intensive. Similar to digital beamforming [59] or MIMO processing [89], the basic formula is as follows. First, the received signal is expressed using an end-to-end transfer function similar to Equation (15) discussed in Section 3.3. Next, the channel matrix is modeled making certain approximations and assumptions. Through careful observation and clever mathematical manipulation, a suitable reward function is formulated.

The resulting beamformer is the result of an optimization process, typically obtained through iterative calculation. The classical approach involves *Kroenecker's singular value decomposition* (SVD) [59] method where the channel matrix is factorized into two unitary matrices and one singular value matrix. The unitary matrices are made of left and right-singular vectors, directly applicable to the beamforming problem.

In addition to this, some knowledge of the channel is required [89]. The required channel state information (CSI) may be obtained using a training scheme of some sort. While efficient channel training is a relevant research topic, it is commonplace in related publications to assume that partial or complete knowledge of the channel

is somehow available at one or both ends of the link. If complete channel knowledge is available, unconstrained rank-1 SVD method minimizes the error probability [28]. Also known as maximal ratio transmission and combining, the method is commonly used as a benchmark when evaluating MIMO precoders.

The method proposed by Nsenga *et al.* [78] follows this approach. Using typical ABF training methods, the channel impulse response (CIR) is estimated for a set of antenna pairs. Next, the SNR is maximized when the first eigenvalue is minimized. This optimization is not, however, straightforward, and the authors resort to maximizing its lower bound instead.

While unconstrained SVD provides best performance in theory it is not the most practical approach because of its complexity as a general algorithm [28]. Moreover, the algorithm cannot be not easily extended to hybrid structures, nor does it leverage the propagation characteristics of the mmW channel. As discussed in Section 3.4, the mmW radio channel is not Rayleigh distributed but sparse and clustered instead. The problem may be further refined based on this observation.

Taking advantage of the sparse nature of millimeter waves, *matching pursuit* (MP) algorithms have gained popularity in the past few years. Originally suggested for mmW beamforming applications by Venkateswaran *et al.* [206], MP is a sparse approximation algorithm for optimal unconstrained precoders. To refine the signal approximation, the greedy algorithm successively picks the best candidate vector from a pool of weight vectors. The vector is added to the active weight vector set found through linear combining to refine the signal approximation. The effect of the active set is then removed from the reward function, and the process is repeated for each transceiver chain.

In literature, one may find numerous mmW hybrid beamformers based on the matching pursuit concept. The proposals are strikingly similar with small differences when it comes to target application, approximations, optimization criteria, and assumptions regarding available CSI and the actual channel model. Some examples of MP algorithms and its variants, such as the orthogonal MP method, are introduced next.

Lin *et al.* [71] employ deliberate noise loading to avoid mismatch problems caused by limited number of training symbols while Kim *et al.* [117] focus on interference suppression. For this, they employ reverse of the interference channel. Although their solution can interface with multiple different analog beamformers and even switched beam antennas, it is not intended for mobile communications.

The proposal by Alkhateeb *et al.* [113] is a good example of assumptions regarding channel knowledge. Both ends of the link are assumed to know their own angles of arrival. Furthermore, assuming reciprocal channel, the AoA correspond to angles of departure (AoD). Using the AoD information, the transmitter sends carefully designed training messages allowing the receiver to estimate properties of different propagation paths. Having this information, the receiver constructs reciprocal channel covariance matrix and uses the MP algorithm to find the precoder weights.

A year later, Alkhateeb *et al.* [76] published an improved hybrid beamformer without making assumptions regarding CSI availability. Furthermore, the proposed system now supports different number of transceiver chains. Interestingly, the approach is based on multiresolution codebooks that are different from common codebooks on two fronts. First, all beams are designed to cover one or more quantized angles at a time. Second, the codebooks combine both analog and digital weighting within the same book.

The algorithm functions by estimating the dominant singular vectors directly from the quantized beam steering directions. As a result, the behavior is somewhat similar to multilevel ABF designs discussed in the previous subsection. First, the channel is estimated by making RSSI measurements with wide beams. The search is then recursively narrowed, concentrating in the direction of maximum RSSI. The multi-receiver architecture is employed to measure multiple directions at one.

Having reached the desired accuracy, adaptive *compressed sensing* tools are leveraged to obtain a channel model of the current path. This model is then used to build both analog and digital weight vectors. These vectors are added to the active set, and their effect on the received signal is erased. This bisection process is then repeated for each transceiver chain similarly to other MP algorithms.

El Ayach *et al.* [134] supplement the MP implementation with basis pursuit functionality they demonstrated [116] two years earlier. Similarly to MP, basis pursuit is a greedy iterative sparse decomposition method. The algorithm exploits the linear transformation property between right-singular SVD vectors and the antenna weight vectors. In their more recent paper [134], the authors also discuss efficient feedback strategies.

The section closes with a real-life implementation. Whereas all the methods discussed above have only been simulations, Yin *et al.* [185] have actually built a prototype of their HBF design. Although the first prototype was presented by Yin *et al.* [185], we discuss the enhanced version with multiuser support reported by Pisek *et al.* [207] a year later. Intended for short-range indoor operation using low-gain antennas, the system operates at 28 GHz operating frequency, and achieves single-stream data rates in excess of 1.5 Gbps with 500 MHz bandwidth.

The access point comprises two phased antenna arrays both of which are divided into two subarrays with their own transceiver. The UE is otherwise similar but has only one antenna array. The baseband implements necessary error correction, modulation, MIMO processing, synchronization, gain control and FFT/IFFT functionality. The system uses multiple OFDM (orthogonal frequency division multiplexing) subcarriers and 16-QAM (quadrature amplitude modulation) with a coding rate of 13/16. With this hardware, the system may be operated in FDMA, in SDMA or real MU-MIMO modes up to two clients.

Some of the OFDM subcarriers are permanently set aside to form a synchronization channel. This channel is not used only for timing and synchronization but is also vital for the beam alignment functionality. The beam pairing process employs successive

approximation based on a multi-level codebook. Within a single level, the search routine is a typical exhaustive search with fast-slow division. The fast AP sends Zadoff-Chu pilots using different beams while the slow UE scans its beams making correlation measurements to determine the best beam pair. The UE then informs the AP of the best beam, and the search continues with the next set of narrower beams. The high-resolution scanning continues every subframe to implement a tracking functionality.

Having found the best TX/RX beams and set-up the link, the MIMO channel is estimated. For this, the transmitter first interleaves a set of pilot bits in the frequency domain before inserting them in every sixth subcarrier. At the receiving end, the channel matrix is estimated using a least squares estimator. Interpolation and linear time-domain estimation is used to find and match the coefficients of each subcarrier. In addition to the subcarrier channel coefficients, the noise variance is estimated using an unused guard band.

8 Summary and conclusions

The number of active subscriptions and mobile broadband data traffic have both increased exponentially during the past decade. Current technologies can support this trend only for a couple of years, and the much needed next-generation systems are indeed currently being outlined. While 5G standardization is still in its infancy, several gigahertz of additional spectral resources are needed globally. Given the abundance of under-utilized spectrum in the V-and E-bands (57 ... 86 GHz), millimeter waves (mmWs) are widely seen as the foremost means of increasing network capacity in dense urban environments.

While 3 ... 5 mm wavelength results in small radio equipment, small wavelength also makes the radio environment extremely challenging. Compared to conventional systems, free-space path loss and atmospheric effects can cause several tens of decibels of extra attenuation. Interaction losses are also multiplied resulting in only a small number of clustered multipath components. While this certainly reduces interference, it also effectively rules out advanced high-rank MIMO processing. While mmW channel modeling is far from complete, it is evident that the links need to be be shorter, and the network devices must be equipped with high-gain antennas.

Consequently, the network will be based on the small cell concept already employed in existing cellular networks. Each macro cell will encompass numerous mmW small cells with a range of just some tens of meters. The small cell APs will have to be installed close to the users, often in suboptimal locations. Wired connectivity might not be available, and wireless techniques are needed for both user access and backhaul. Given the small hop length and obstructed street-level environment, the APs need to be chained to be able to cover the entire macro cell. Moreover, since blockages in the LOS path quickly render the links unusable, the APs are further connected as a mesh network.

Given the high-gain antennas with pencil-like beams, careful link alignment is required. Furthermore, due to the sheer number of devices, any tedious on-site configuration such as manual antenna pointing is economically impossible. Moreover, the installations are prone to small yet significant sway and twist that can cause severe degradation of link quality. The network and the devices should thus exhibit high level of automation, and not only form the links automatically but also actively compensate for any deflections. Efficient beam alignment studied in this thesis is one of the key enablers for mmW mobile communications.

In terms of antenna hardware, electronic beam steering is typically based on array technology in one way or another. There are two main approaches, namely beam switching and beamforming. Beam switched antennas such as integrated lens antennas, Butler matrices or Rotman lenses, are steered by switching between multiple feed ports. While beam switching is simple and commercially attractive, beamforming is ultimately preferred thanks to its versatility. By controlling the phase and possibly amplitude of individual antenna elements, a beamformer may optimize its radiation pattern to fit the link requirements. More advanced processing such

as true P2MP and/or MIMO operation is possible if the BF comprises of multiple transceivers. In practice, however, not only is the number of TRX chains is typically very limited but also the antenna weights are typically predetermined and taken from a codebook.

While all beam alignment can be thought of channel estimation, the available hardware essentially dictates the actual functionality. Both ends of the link need to be aligned simultaneously, meaning antenna alignment concerns finding the best beam pair. The straightforward solution is to test all possible beam pairs and use the best beam for subsequent communication. While this exhaustive searching forms the basis of all beam steering, it is also very inefficient.

In single-receiver systems, the alignment effort may be reduced in three ways. First, instead of finding the optimum pair, the search could be preempted after finding the first pair of sufficient quality. Second, the effective search area may be reduced though hierarchical searching. While this approach is employed in 60 GHz WLAN and WPAN systems, it requires variable beamwidth. The third approach employs the principles of active search where all measurement steps contribute to the final solution. While the aforementioned methods decrease the testing effort, they also require efficient feedback which is a major design challenge especially in half-duplex systems. In this cases, the nodes are best equipped with a secondary control link operating in a different frequency band.

The aforementioned approaches work also with multi-transceiver systems. Based on advanced DSP, digital beamformers may further exploit the orthogonality of signal and noise subspaces for improved accuracy. In hybrid beamforming, DSP is typically used for MIMO processing while beams are still generated using analog weights. Typically these weights are tuned iteratively based on the concept of matching pursuit – an optimization algorithm exploiting the sparse nature of the mmW channel. Even though some of the reported approaches show great promise, they have only been verified using simulations based on incomplete mmW channel models.

There are, however, bigger problems. Virtually all existing mmW solutions target short-range indoor communications and consider only a single (access) link between two independent devices. As the requirements, goals and underlying constraints are entirely different between the applications, aforementioned approaches cannot be directly applied in mmW cellular communications – not in mobile access nor backhauling. To address these concerns, this thesis considers both the physics of the situation and the surrounding system as a whole.

The work investigates a single mesh network comprising of numerous wireless mmW APs and a macro BS doubling as a controller and a gateway. Targeting for relaxed hardware requirements, the nodes are equipped with only a single mmW TRX, and both duplexing and multiple-access are based on time-division. Furthermore, the network is to be built in an incremental fashion with minimal effort or intervention simply by mounting the devices in suitable locations – everything else is automatic.

The expansion process is similar to installing a new router, and the devices are completely unaware of the existing network. The devices are initially also unsynchronized and incapable of connecting to the mmW mesh. A node must thus gain synchronization, download latest configuration and align its beams before it can provide any service. By purposeful organization and exploitation of the existing communications infrastructure, these tasks may be carried out with minimal impact on the existing network.

The process is outlined as follows. Equipped with a secondary low-rate control connection, a GNSS receiver and an orientation sensor, the node first connects to the macro network not only to synchronize and configure itself but also to authenticate and declare its existence and location. This triggers a network-wide update where both the network schedule and the routing tables are changed to reflect a new node. The node with the least impact on current operation is scheduled to pair with the new device. This information is relayed to the new node which is then ready for pairing. During the specified training slot, the nodes exploit location information to align their beams quickly before completing the link setup.

Proceeding in this way has numerous benefits. First, the network is responsive, and the new node may begin serving clients posthaste. Second, all unnecessary scanning is eliminated. Third, time consuming synchronization and other control tasks are offloaded to a secondary link. Furthermore, the nodes may use the link also for feedback and the nodes stay connected even though the mmW link is dropped. Moreover, to combat outages and improve service, the network will seek to form alternate paths and/or fine tune their alignment during quiet hours.

In addition to building the network by mapping the nearby devices, automatic beam alignment is needed to counter the side effects of suboptimal installations prone to vibration and sway. The misalignment phenomenon was modeled using an endpoint attached to a rigid pole pivoting from its root. Due to leverage, pivoting causes both rotation and movement. Out of the two, the link is more susceptible to the rotation due to extremely narrow beamwidths. Luckily, the effects of rotation are confined to the rotating end, meaning it can be self-corrected without any training using an orientation sensor.

While sensor accuracy typically leaves room for improvement, they can regardless be used to estimate the direction of movement to concentrate training effort in the estimated direction. The search area may be further reduced if limited artificial intelligence is introduced to the system. Primarily due to wind, the deflection behavior is mostly periodic, meaning it can be characterized and even predicted by applying the principles of pattern recognition. In case of unpredictable conditions, however, coarsely aligned wide beams may prove to be the best alternative.

Author's vision of a mmW small cell system exploits the existing network infrastructure, location and orientation information, and even artificial intelligence to provide ultra high QoS with minimal cost and overhead. At first, the system will based on beam switching, later on analog beamforming, and ultimately on hybrid

beamforming. Such a system will be capable of multi-Gbps data rates and true P2MP/MIMO operation using two orthogonal polarizations. However, much research effort must go into beamforming systems, mmW channel modeling, alignment algorithms and efficient signaling before commercial deployment is possible. Luckily, the topic has gained world-wide attention, and researchers are rising to the challenge.

It can thus be concluded that the work is well under way, and commercial deployment of next-generation systems in the early 2020's seems realistic. Through 5G networking, billions of devices and users can exchange data with high throughput, low latency and with extreme reliability. Truly ubiquitous computing anytime and everywhere opens entirely another frontier of applications with infinite possibilities we cannot event begin to imagine. Immersed in a connected world, we are indeed on the verge of a new revolution.

References

[1] Accenture, "The 2012 Accenture consumer electronics products and services usage report," Dublin, Ireland, 2012. [Online]. Available: http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture_EHT_Research_2012_Consumer_Technology_Report.pdf

[2] ——, "The 2013 Accenture consumer electronics products and services usage report," Dublin, Ireland, 2013. [Online]. Available: http://www.accenture.com/SiteCollectionDocuments/PDF/2013-Accenture-Consumer-Electronics-Products-and-Services-Usage-Report.pdf

[3] ——, "Engaging the digital consumer in the new connected world," Dublin, Ireland, 2014. [Online]. Available: http://www.accenture.com/SiteCollectionDocuments/us-en/engaging-digital-consumer-new-connected-world.pdf

[4] ——, "Accenture digital consumer tech survey 2014," Dublin, Ireland, 2014. [Online]. Available: http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Digital-Consumer-Tech-Survey-2014

[5] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2014–2019," San Jose, CA, USA, 2015. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[6] Ericsson, "Ericsson mobility report – On the pulse of the networked society," Stockholm, Sweden, 2014. [Online]. Available: http://www.ericsson.com/res/docs/2014/ericsson-mobility-report-november-2014.pdf

[7] Nokia, "Looking ahead to 5G," Espoo, Finland, 2014. [Online]. Available: http://networks.nokia.com/sites/default/files/document/5g_white_paper_0.pdf

[8] METIS, "Deliverable D1.1 – Scenarios, requirements and KPIs for 5G mobile and wireless systems," Stockholm, Sweden, 2013. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.1_v1.pdf

[9] ——, "Deliverable D1.5 – Updated scenarios, requirements and KPIs for 5G mobile and wireless system with recommendations for future investigations," Stockholm, Sweden, 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.5_v1.pdf

[10] NGMN Alliance, "5G white paper," Frankfurt, Germany, 2015. [Online]. Available: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf

[11] Nokia, "5G radio access – System design aspects," Espoo, Finland, 2015. [Online]. Available: http://networks.nokia.com/sites/default/files/document/nokia_5g_radio_access_white_paper.pdf

[12] 4G Americas, "Meeting the 1000x challenge," Bellevue, WA, USA, 2013. [Online]. Available: http://www.4gamericas.org/files/3614/0758/8793/4G_Americas_Meeting_the_1000x_Challenge_Condensed_May_2014_FINAL.pdf

[13] ——, "Beyond LTE: Enabling the mobile broadband explosion," Bellevue, WA, USA, 2014. [Online]. Available: http://www.4gamericas.org/files/7514/1021/4070/Beyond_LTE_Enabling_Mobile_Broadband_Explosion_August_2014x.pdf

[14] Nokia, "5G use cases and requirements," Espoo, Finland, 2014. [Online]. Available: http://networks.nokia.com/sites/default/files/document/5g_requirements_white_paper.pdf

[15] ——, "Network architecture for the 5G era," Espoo, Finland, 2015. [Online]. Available: http://networks.nokia.com/sites/default/files/document/nokia_5g_architecture_white_paper.pdf

[16] METIS, "Deliverable D6.6 – Final report on the METIS 5G system concept and technology roadmap," Stockholm, Sweden, 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D6.6_v1.pdf

[17] Huawei, "5G: A technology vision," Shezhen, China, 2013. [Online]. Available: http://www.huawei.com/5gwhitepaper/

[18] Alcatel-Lucent, "5G is coming," Boulogne-Billancourt, France, 2015. [Online]. Available: http://www2.alcatel-lucent.com/landing/5g/

[19] METIS, "Deliverable D8.4 – METIS final project report," Stockholm, Sweden, 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D8.4_v1.pdf

[20] 4G Americas, "4G Americas' recomendations on 5G – Requirements and solutions," Bellevue, WA, USA, 2014. [Online]. Available: http://www.4gamericas.org/files/2714/1471/2645/4G_Americas_Recommendations_on_5G_Requirements_and_Solutions_10_14_2014-FINALx.pdf

[21] 5G Forum Korea, "5G vision, requirements, and enabling technologies," Seoul, South Korea, 2015. [Online]. Available: http://media.wix.com/ugd/9b2680_c0db0a0488044e9ba8ead9c8e1c3eb53.pdf

[22] IMT-2020 (5G) Promotion Group, "5G vision and requirements," Beijing, China, 2015. [Online]. Available: http://www.imt-2020.cn/en/documents/download/3

[23] GSMA Intelligence, "Understanding 5G: Perspectives on future technological advancements in mobile," London, UK, 2014. [Online]. Available: https://gsmaintelligence.com/research/?file=141208-5g.pdf&download

[24] 5G PPP, "5G vision," Heidelberg, Germany, 2015. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf

[25] VTT, "5G test network," 2015. [Online]. Available: http://www.5gtn.fi/

[26] Industrie 4.0 Working Group, "Recommendations for implementing the strategic initiative INDUSTRIE 4.0," Berlin, Germany, 2013. [Online]. Available: http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonderseiten/Industrie_4.0/Final_report__Industrie_4.0_accessible.pdf

[27] Nokia Siemens Networks, "2020: Beyond 4G – Radio evolution for the gigabit experience," Espoo, Finland, 2011. [Online]. Available: http://gr.networks.nokia.com/system/files/document/nokia_siemens_networks_beyond_4g_white_paper_online_20082011_0.pdf

[28] T. Kim, J. Park, J. Y. Seol, S. Jeong, J. Cho, and W. Roh, "Tens of Gbps support with mmWave beamforming systems for next generation communications," in *2013 IEEE Global Communications Conference*, Atlanta, GA, USA, Dec. 9–13, 2013, pp. 3685–3690. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6831646

[29] K. Seppänen, T. Suihko, and P. Wainio, "Self-optimizing last hop backhaul for 5G," *Unpublished*, 2015.

[30] NGMN Alliance, "Small cell backhaul requirements," Frankfurt, Germany, 2012. [Online]. Available: https://www.ngmn.org/uploads/media/NGMN_Whitepaper_Small_Cell_Backhaul_Requirements.pdf

[31] J. Wannstrom and K. Mallinson, "HetNet/Small cells," 2012. [Online]. Available: http://www.3gpp.org/hetnet

[32] Small Cell Forum, "Backhaul technologies for small cells: Use cases, requirements and solutions," Dursley, UK, 2013. [Online]. Available: http://www.scf.io/en/get_email.php?doc=049

[33] K. Sakaguchi, G. K. Tran, H. Shimodaira, S. Nanba, T. Sakurai, K. Takinami, I. Siaud, E. C. Strinati, A. Capone, I. Karls, R. Arefi, and T. Haustein, "Millimeter-wave evolution for 5G cellular networks," *IEICE Transactions on Communications*, vol. E98-B, no. 3, pp. 388–402, Mar. 2015. [Online]. Available: http://search.ieice.org/bin/summary.php?id=e98-b_3_388

[34] MiWaveS, "Beyond 2020 heterogenous wireless network with millimeter wave small cell access and backhauling," *Unpublished*, 2013.

[35] D. Chen, J. Schuler, P. Wainio, and J. Salmelin, "5G self-optimizing wireless mesh backhaul," *Unpublished*, 2015.

[36] K. Seppänen, J. Kilpi, J. Paananen, T. Suihko, P. Wainio, and J. Kapanen, "Resilient multipath routing for static millimeter wave wireless mesh networks," *Unpublished*, 2015.

[37] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, Feb. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6736750

[38] N. Guo, R. C. Qiu, S. S. Mo, and K. Takahashi, "60-GHz millimeter-wave radio: Principle, technology, and new results," *EURASIP Journal on Wireless Communications and Networking*, Dec. 2007. [Online]. Available: http://jwcn.eurasipjournals.com/content/2007/1/068253

[39] P. Adhikari, "Understanding millimeter wave wireless communication," San Diego, CA, USA, 2008. [Online]. Available: http://www.nocsta.com/pages/pdffiles/L1104-WP_UnderstandingMMWComcopy.pdf

[40] IEEE Computer Society, "IEEE standard for information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements, part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specification," New York, NY, USA, 2012. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11ad-2012.pdf

[41] ——, "IEEE standard for information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements, part 15.3: Wireless medium access control (MAC) and physical layer (PHY) specification," New York, NY, USA, 2009. [Online]. Available: http://standards.ieee.org/getieee802/download/802.15.3c-2009.pdf

[42] ECMA International, "Standard ECMA-387: High rate 60 GHz PHY, MAC and PALs," Geneva, Switzerland, 2010. [Online]. Available: http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-387.pdf

[43] WirelessHD Consortium, "WirelessHD specification v.1.1 overview," Morgan Hill, CA, USA, 2010. [Online]. Available: http://www.wirelesshd.org/pdfs/WirelessHD-Specification-Overview-v1.1May2010.pdf

[44] R. Kalimulin, A. Artemenko, J. Putkonen, and J. Salmelin, "Impact of mounting structures twists and sways on point-to-point millimeter-wave backhaul links," *Unpublished*, 2015.

[45] T. S. Rappaport, R. Mayzus, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6515173

[46] M. Kyrö, "Radio wave propagation and antennas for millimeter-wave communications," D. Sc. (Tech.) dissertation, Aalto University, 2012. [Online]. Available: http://urn.fi/URN:ISBN:978-952-60-4951-9

[47] T. A. Thomas, H. C. Nguyen, G. R. MacCartney, and T. S. Rappaport, "3D mmWave channel model proposal," in *2014 IEEE Vehicular Technology Conference*, Vancouver, Canada, Sep. 14–17, 2014, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6965800

[48] METIS, "Deliverable D1.4 – METIS channel models," Stockholm, Sweden, 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.4_v2.pdf

[49] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6732923

[50] VIBRU, "Deliverable D2.2: Defining statistics required for beamsteering algorithm development," *Unpublished*, 2014.

[51] A. Drozdy and J. Kapanen, "Millimeter wave in-band mobile access and backhaul for 5G," *Unpublished*, 2015.

[52] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6600706

[53] D. Pozar, *Microwave Engineering*, 4th ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2012. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-EHEP002016.html

[54] C. Balanis, *Antenna Theory – Analysis and Design*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2005. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-047166782X.html

[55] J. Volakis, *Antenna Engineering Handbook*, 4th ed. New York, NY, USA: McGraw-Hill, 2007. [Online]. Available: http://www.mhprofessional.com/product.php?isbn=0071475745

[56] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2013. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-EHEP002012.html

[57] C. Balanis and P. Ioannides, *Introduction to Smart Antennas*. San Rafael, CA, USA: Morgan & Claypool, 2007. [Online]. Available: http://www.morganclaypool.com/doi/abs/10.2200/S00079ED1V01Y200612ANT005

[58] R. J. Mailloux, *Phased Array Antenna Handbook*, 2nd ed. Norwood, MA, USA: Artech House, 2005. [Online]. Available: http://www.artechhouse.com/Main/Books/Phased-Array-Antenna-Handbook-Second-Edition-982.aspx

[59] W. Liu and S. Weiss, *Wideband Beamforming: Concepts and Techniques*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2010. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470713925.html

[60] D. Parker and D. Zimmermann, "Phased arrays – Part II: Implementations, applications, and future trends," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 688–698, Mar. 2002. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=989954

[61] A. J. Fenn, D. H. Temme, W. P. Delaney, and W. E. Courtney, "The development of phased-array radar technology," *Lincoln Laboratory Journal*, vol. 12, no. 2, pp. 321–340, Feb. 2000. [Online]. Available: https://www.ll.mit.edu/publications/journal/pdf/vol12_no2/12_2devphasedarray.pdf

[62] J. Kraus, *Antennas*, 2nd ed. New Delhi, India: Tata McGraw-Hill, 2001.

[63] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Multilevel millimeter wave beamforming for wireless backhaul," in *2011 IEEE Global Telecommunications Conference*, Houston, TX, USA, Dec. 5–9, 2011, pp. 253–257. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6162448

[64] R. Kinsey, "Phased array beam spoiling technique," in *1997 IEEE Antennas and Propagation Society International Symposium*, vol. 2, Montreal, Canada, Jul. 13–18, 1997, pp. 698–701. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=631557

[65] L. Zhou and Y. Ohashi, "Efficient codebook-based MIMO beamforming for millimeter-wave WLANs," in *2012 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sydney, Australia, Sep. 9–12, 2012,

pp. 1885–1889. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6362659

[66] B. Li, Z. Zhou, W. Zou, X. Sun, and G. Du, "On the efficient beam-forming training for 60 GHz wireless personal area networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 504–515, Feb. 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6399486

[67] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, "MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?" *IEEE Communications Magazine*, vol. 52, no. 12, pp. 110–121, Dec. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6979962

[68] W. Feng, Z. Xiao, D. Jin, and L. Zeng, "Circular-antenna-array-based codebook design and training method for 60GHz beamforming," in *2013 IEEE Wireless Communications and Networking Conference*, Shanghai, China, Apr. 7–10, 2013, pp. 4140–4145. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6555241

[69] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1390–1399, Aug. 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5262295

[70] C. Wu, H. Zhang, X. Cui, and G. Zhang, "Codebook beam switching based relay scheme for 60 GHz anti-blockage communication," in *2013 International Conference on Emerging Intelligent Data and Web Technologies*, Xi'an, China, Sep. 9–11, 2013, pp. 197–203. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6631617

[71] Z. Lin, X. Peng, and F. Chin, "Enhanced beamforming for 60 GHz OFDM system with co-channel interference mitigation," in *2011 IEEE International Conference on Ultra-Wideband*, Bologna, Italy, Sep. 14–16, 2011, pp. 29–33. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6058850

[72] MiWaveS, "Deliverable D2.2 – Automatic antenna beam steering and beamforming," *Unpublished*, 2015.

[73] P. Pal and D. Mondal, "Adaptive antenna array pattern synthesis for suppressed sidelobe level and controlled null using genetic algorithm," in *2014*

International Conference for Convergence for Technology, Pune, India, Apr. 6–8, 2014, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7092156

[74] A. Tamminen, J. Ala-Laurinaho, S. Mäkelä, A. V. Räisänen, D. G. Martins, J. Hakli, P. Koivisto, P. Rantakari, J. Säily, R. Tuovinen, and A. Luukanen, "Millimeter-wave reflectarray for beam-steering applications," in *2012 European Microwave Integrated Circuits Conference*, Amsterdam, The Netherlands, Sep. 7–8, 2012, pp. 219–222. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6483775

[75] S.-K. Yong, P. Xia, and A. Valdes-Garcia, *60 GHz Technology for Gbps WLAN and WPAN: From Theory to Practice*. New York, NY, USA: John Wiley & Sons, Inc., 2010. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470747706.html

[76] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6847111

[77] X. Huang, Y. Guo, and J. Bunton, "A hybrid adaptive antenna array," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1770–1779, May 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5463231

[78] J. Nsenga, A. Bourdoux, and F. Horlin, "Mixed analog/digital beamforming for 60 GHz MIMO frequency selective channels," in *2010 IEEE International Conference on Communications*, Cape Town, South Africa, May 23–27, 2010, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5502689

[79] Y. J. Guo, X. Huang, and V. Dyadyuk, "A hybrid adaptive antenna array for long-range mm-wave communications," *IEEE Antennas and Propagation Magazine*, vol. 54, no. 2, pp. 271–282, Apr. 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6230773

[80] S. Han, C.-L. I, Z. Xu, and C. Rowel, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 186 – 194, Jan. 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7010533

[81] L. Kong, "Energy-efficient 60 GHz phased-array design for multi-Gbps communication systems," Ph. D. dissertation, University of

California at Berkeley, Berkeley, CA, USA, 2014. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-191.pdf

[82] J. P. A. Pérez, S. C. Pueyo, and B. C. López, *Automatic Gain Control: Techniques and Architectures for RF Receivers*. New York, NY, USA: Springer Publishing, 2011. [Online]. Available: http://www.springer.com/us/book/9781461401667

[83] S. R. Bullock, *Transceiver and System Design for Digital Communications*, 4th ed. Edison, NJ, USA: SciTech Publishing, Inc., 2014. [Online]. Available: http://digital-library.theiet.org/content/books/cs/sbcs504e

[84] M. Golio, *The RF and Microwave Handbook*. Boca Raton, FL, USA: CRC Press LLC, 2001. [Online]. Available: https://www.crcpress.com/The-RF-and-Microwave-Handbook/Golio/9781420036763

[85] K. Chang, *RF and Microwave Wireless Systems*. New York, NY, USA: John Wiley & Sons, Inc., 2000. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471351997.html

[86] A. Ghasemi, A. Abedi, and F. Ghasemi, *Propagation Engineering in Wireless Communications*. New York, NY, USA: Springer Publishing, 2012. [Online]. Available: http://www.springer.com/us/book/9781461453130

[87] T. Manning, *Microwave Radio Transmission Design Guide*, 2nd ed. Norwood, MA, USA: Artech House, 2009. [Online]. Available: http://www.artechhouse.com/Main/books.aspx?iid=978-1-59693-456-6

[88] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Singapore City, Singapore: John Wiley & Sons (Asia) Pte Ltd., 2010. [Online]. Available: www.wiley.com/go/chomimo

[89] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge, UK: Cambridge University Press, 2010. [Online]. Available: http://www.cambridge.org/us/academic/subjects/engineering/wireless-communications/mimo-wireless-communications

[90] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, 2005. [Online]. Available: http://www.eecs.berkeley.edu/~dtse/book.html

[91] O. Jo, W. Hong, S. Choi, S. Chang, C. Kweon, J. Oh, and K. Cheun, "Holistic design considerations for environmentally adaptive 60 GHz beamforming technology," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 30–38, Nov. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6957140

[92] X. Huang and Y. J. Guo, "Frequency-domain AoA estimation and beamforming with wideband hybrid arrays," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2543–2553, Aug. 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5934684

[93] A. Karttunen, "Millimetre and submillimetre wave antenna design using ray tracing," D. Sc. (Tech.) dissertation, Aalto University, 2013. [Online]. Available: http://urn.fi/URN:ISBN:978-952-60-5293-9

[94] K. R. Jha and G. Singh, *Terahertz Planar Antennas for Next Generation Communication*. New York, NY, USA: Springer Publishing, 2014. [Online]. Available: http://www.springer.com/us/book/9783319023403

[95] Radio Gigabit, "Mayak-1000 – 1 Gbps 60 GHz reference link," 2014. [Online]. Available: http://www.radiogigabit.com/mayak-1000/

[96] A. Artemenko, A. Mozharovskiy, A. Sevastyanov, V. Ssorin, and R. Maslennikov, "High gain lens antennas for 71–86 GHz point-to-point applications," in *2013 European Microwave Conference*, Nuremberg, Germany, Oct. 6–11, 2013, pp. 361–364. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6686666

[97] D. H. Archer and M. J. Maybell, "Rotman lens development history at Raytheon Electronic Warfare Systems 1967-1995," in *2005 IEEE Antennas and Propagation Society International Symposium*, vol. 2B, Washington, DC, USA, Jul. 3–8, 2005, pp. 31–34. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1551927

[98] M. Pokorný, "Rotman lens desing with HFSS link," 2015. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/50490-rotman-lens-desing-with-hfss-link

[99] H. Lehpamer, *Microwave Transmission Networks: Planning, Design, and Deployment*, 2nd ed. New York, NY, USA: McGraw-Hill, 2010.

[100] A. Goldsmith, *Wireless Communications*. Cambridge, UK: Cambridge University Press, 2005. [Online]. Available: http://www.cambridge.org/us/academic/subjects/engineering/wireless-communications/wireless-communications

[101] ITU-R, "Recommendation P.530-16: Propagation data and prediction methods required for the design of terrestrial line-of-sight systems," Geneva, Switzerland, 2015. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.530-16-201507-I!!PDF-E.pdf

[102] ——, "Recommendation P.676-10: Attenuation by atmospheric gases," Geneva, Switzerland, 2013. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.676-10-201309-I!!PDF-E.pdf

[103] ——, "Recommendation P.837-6: Characteristics of precipitation for propagation modeling," Geneva, Switzerland, 2012. [Online]. Available: http://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.837-6-201202-I!!PDF-E.pdf

[104] ——, "Recommendation PN.837-1: Characteristics of precipitation for propagation modeling," Geneva, Switzerland, 1994. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.837-1-199408-S!!PDF-E.pdf

[105] ——, "Recommendation P.838-3: Specific attenuation model for rain for use in prediction methods," Geneva, Switzerland, 2005. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.838-3-200503-I!!PDF-E.pdf

[106] ——, "Recommendation P.840-6: Attenuation due to clouds and fog," Geneva, Switzerland, 2013. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.840-6-201309-I!!PDF-E.pdf

[107] H. V. Le, T. Hirano, J. Hirokawa, and M. Ando, "Site diversity performance of millimeter wave wireless networks against localized rain," in *2014 European Conference on Antennas and Propagation*, The Hague, The Netherlands, Apr. 6–11, 2014, pp. 3477–3481. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6902578

[108] N. Sheng, C. Liao, W. Lin, Q. Zhang, and R. Bai, "Modeling of millimeter-wave propagation in rain based on parabolic equation method," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 3–6, Jan. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6680617

[109] Federal Communications Commission, "FCC 13-112: Revision of part 15 of the Commission's rules regarding operation in the 57-64 GHz band," Washington, DC, USA, 2013. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/DA-13-2403A1.pdf

[110] K. Haneda, C. Gustafson, and S. Wyne, "60 GHz spatial radio transmission: Multiplexing or beamforming?" *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 11, pp. 5735–5743, Nov. 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6583281

[111] METIS, "Deliverable D6.1 – Simulation guidelines," Stockholm, Sweden, 2013. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D6.1_v1.pdf

[112] A.-S. K. Pathan, M. M. Monowar, and S. Khan, *Simulation Technologies in Networking and Communications: Selecting the Best Tool for the Test*. Boca Raton, FL, USA: CRC Press, 2015. [Online]. Available: https://www.crcpress.com/Simulation-Technologies-in-Networking-and-Communications-Selecting-the/Pathan-Monowar-Khan/9781482225495

[113] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *2013 Information Theory and Applications Workshop*, San Diego, CA, USA, Feb. 10–15, 2013, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6522603

[114] S. G. Larew, D. J. Love, T. A. Thomas, and A. Ghosh, "Millimeter wave beam-alignment for dual-polarized outdoor MIMO systems," in *2013 IEEE Global Communications Conference*, Atlanta, GA, USA, Dec. 9–13, 2013, pp. 356–361. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6825013

[115] J. Song, S. G. Larew, D. J. Love, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for multiuser dual-polarized MIMO systems," in *2013 IEEE Global Conference on Signal and Information Processing*, Austin, TX, USA, Dec. 3–5, 2013, pp. 719–722. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6736992

[116] O. El Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *2012 IEEE International Conference on Communications*, Ottawa, Canada, Jun. 10–15, 2012, pp. 3724–3729. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6363634

[117] M. Kim and Y. H. Lee, "MSE-based hybrid RF/baseband processing for millimeter wave communication systems in MIMO interference channels," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2714–2720, Jun. 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6874567

[118] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2007.

[119] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-wave multipath clustering and channel modeling," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1445–1455, Mar. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6691924

[120] Q. C. Li, G. Wu, and T. S. Rappaport, "Channel model for millimeter-wave communications based on geometry statistics," in *2014 IEEE Global Communications Conference*, Austin, TX, USA, Dec. 8–12, 2014, pp. 427–432. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7063469

[121] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter-wave propagation measurements

and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1850–1859, Apr. 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6387266

[122] G. R. MacCartney and T. S. Rappaport, "Path loss models for 5G millimeter wave propagation channels in urban microcells," in *2013 IEEE Global Communications Conference*, Atlanta, GA, USA, Dec. 9–13, 2013, pp. 3948–3953. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6831690

[123] J. S. Seybold, *Introduction to RF Propagation*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2005. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471655961.html

[124] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications," in *2012 IEEE International Conference on Communications*, Ottawa, Canada, Jun. 10–15, 2012, pp. 4568–4573. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6363891

[125] H. Zhang, S. Venkateswaran, and U. Madhow, "Channel modeling and MIMO capacity for outdoor millimeter wave links," in *2010 IEEE Wireless Communication and Networking Conference*, Sydney, Australia, Apr. 18–21, 2010, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5506714

[126] S. Mota, M. Garcia, A. Rocha, and F. Perez-Fontan, "Clustering of the multipath radio channel parameters," in *2011 European Conference on Antennas and Propagation*, Rome, Italy, Apr. 11–15, 2011, pp. 3232–3236. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5782271

[127] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6834753

[128] J. Medbo, K. Börner, K. Haneda, V. Hovinen, T. Imai, J. Järveläinen, T. Jämsä, A. Karttunen, K. Kusume, J. Kyröläinen, P. Kyösti, J. Meinilä, V. Nurmela, L. Raschkowski, A. Roivainen, and J. Ylitalo, "Channel modelling for the fifth generation mobile communications," in *2014 European Conference on Antennas and Propagation*, The Hague, The Netherlands, Apr. 6–11, 2014, pp. 219–223. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6901730

[129] M. Kyrö, S. Ranvier, V. Kolmonen, K. Haneda, and P. Vainikainen, "Long range wideband channel measurements at 81–86 GHz frequency range," in *2010 European Conference on Antennas and Propagation*, Barcelona, Spain, Apr. 12–16, 2010, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5505183

[130] T. S. Rappaport, Y. Qiao, J. I. Tamir, J. N. Murdock, and E. Ben-Dor, "Cellular broadband millimeter wave propagation and angle of arrival for adaptive beam steering systems," in *2012 IEEE Radio and Wireless Symposium*, Santa Clara, CA, USA, Jan. 15–18, 2012, pp. 151–154. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6175397

[131] A. Maltsev, E. Perahia, R. Maslennikov, A. Sevastyanov, A. Lomayev, and A. Khoryaev, "Impact of polarization characteristics on 60-GHz indoor radio communication systems," *IEEE Antennas and Wireless Propagation Letters*, vol. 9, pp. 413–416, Apr. 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5451154

[132] S. Wyne, K. Haneda, S. Ranvier, F. Tufvesson, and A. F. Molisch, "Beamforming effects on measured mm-wave channel characteristics," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3553–3559, Nov. 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6015598

[133] A. V. Räisänen, J. Ala-Laurinaho, K. Haneda, J. Järveläinen, A. Karttunen, M. Kyrö, V. Semkin, A. Lamminen, and J. Säily, "Studies on E-band antennas and propagation," in *2013 Loughborough Antennas & Propagation Conference*, Loughborough, UK, Nov. 11–12, 2013, pp. 176–180. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6711876

[134] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6717211

[135] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications," in *2014 IEEE Global Communications Conference*, Austin, TX, USA, Dec. 8–12, 2014, pp. 3326–3331. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7037320

[136] S. Sun and T. S. Rappaport, "Wideband mmWave channels: Implications for design and implementation of adaptive beam antennas," in *2014 IEEE MTT-S International Microwave Symposium*, Tampa Bay, FL, USA, Jun. 1–6,

2014, pp. 1–4. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6848596

[137] T. Rappaport, W. Roh, and K. Cheun, "Mobile's millimeter-wave makeover," *IEEE Spectrum*, vol. 51, no. 9, pp. 34–58, Sep. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6882985

[138] S. Nie, G. R. MacCartney, S. Sun, and T. S. Rappaport, "28 GHz and 73 GHz signal outage study for millimeter wave cellular and backhaul communications," in *2014 IEEE International Conference on Communications*, Sydney, Australia, Jun. 10–14, 2014, pp. 4856–4861. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6884089

[139] P. Xia, S.-K. Yong, J. Oh, and C. Ngo, "A practical SDMA protocol for 60 GHz millimeter wave communications," in *2008 Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Oct. 26–29, 2008, pp. 2019–2023. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5074786

[140] J. Jervase, "Comparison of cross polar discrimination in circularly and linearly polarized reflector antennas with surface imperfections," in *1993 International Conference on Antennas and Propagation*, Edinburgh, UK, Mar. 30 – Apr. 2, 1993, pp. 554–558 vol. 1. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=224839

[141] E. Hecht, *Optics*, 4th ed. Boston, MA, USA: Addison-Wesley, 2002. [Online]. Available: http://www.pearsonhighered.com/educator/product/Optics/9780805385663.page

[142] ETSI, "TR 125 996 V12.0.0 (2014-09) – Universal Mobile Telecommunications System (UMTS) – Spatial channel model for multiple input multiple output (MIMO) simulations," Valbonne, France, 2007. [Online]. Available: http://www.etsi.org/deliver/etsi_tr/125900_125999/125996/12.00.00_60/tr_125996v120000p.pdf

[143] 3GPP, "3GPP TR 36.873 V12.2.0 (2015-06) – Study on 3D channel model for LTE," Valbonne, France, 2015. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/36_series/36.873/36873-c20.zip

[144] Information Society Technologies, "IST-2003-507581 WINNER deliverable D5.4 v.1.4 – Final report on link level and system level channel models," Munich, Germany, 2005. [Online]. Available: http://www.ist-winner.org/DeliverableDocuments/D5.4.pdf

[145] ——, "IST-4-027756 WINNER II deliverable D1.1.2 v.1.1 – WINNER II channel models – Part I: Channel models," Munich, Germany, 2007. [Online]. Available: http://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf

[146] CELTIC, "CELTIC / CP5-026 deliverable D5.3 – WINNER+ final channel models," Heidelberg, Germany, 2010. [Online]. Available: http://projects.celtic-initiative.org/winner+/WINNER+Deliverables/D5.3_v1.0.pdf

[147] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 92–99, Dec. 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6393523

[148] ITU-R, "Report M.2135: Guidelines for evaluation of radio interface technologies for IMT-Advanced," Geneva, Switzerland, 2009. [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2135-1-2009-PDF-E.pdf

[149] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, Feb. 1987. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1146527

[150] IEEE P802.11, "Channel models for 60 GHz WLAN systems," New York, NY, USA, 2010. [Online]. Available: https://mentor.ieee.org/802.11/dcn/09/11-09-0334-08-00ad-channel-models-for-60-ghz-wlan-systems.doc

[151] S. Loredo, A. Rodriguez-Alonso, and R. Torres, "Indoor MIMO channel modeling by rigorous GO/UTD-based ray tracing," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 2, pp. 680–692, Mar. 2008. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4357313

[152] K. Taylor, B. Walker, R. Graham, and J. Matusiak, "Using site-specific, ray-tracing channel models to control wireless testbeds," in *2013 IEEE Military Communications Conference*, San Diego, CA, USA, Nov. 18–20, 2013, pp. 317–322. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6735642

[153] H. Asif, B. Honary, and H. Ahmed, "Multiple-input multiple-output ultra-wide band channel modelling method based on ray tracing," *IET Communications*, vol. 6, no. 10, pp. 1195–1204, Jun. 2012. [Online]. Available: http://digital-library.theiet.org/content/journals/10.1049/iet-com.2011.0265

[154] J. McKown and R. Hamilton, "Ray tracing as a design tool for radio networks," *IEEE Network*, vol. 5, no. 6, pp. 27–30, Nov. 1991. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=103807

[155] O. Stäbler and R. Hoppe, "MIMO channel capacity computed with 3D ray tracing model," in *2009 European Conference on Antennas and Propagation*,

Berlin, Germany, Mar. 23–27, 2009, pp. 2271–2275. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5068069

[156] T. Kurner and M. Jacob, "Application of ray tracing to derive channel models for future multi-gigabit systems," in *2009 International Conference on Electromagnetics in Advanced Applications*, Torino, Italy, Sep. 14–18, 2009, pp. 517–520. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5297379

[157] E. Ben-Dor, T. S. Rappaport, and S. J. Lauffenburger, "Millimeter-wave 60 GHz outdoor and vehicle AOA propagation measurements using a broadband channel sounder," in *2011 IEEE Global Telecommunications Conference*, Houston, TX, USA, Dec. 5–9, 2011, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6133581

[158] A. Muller, "Monte-Carlo multipath simulation of ray tracing channel models," in *1994 IEEE Global Communications Conference*, vol. 3, San Francisco, CA, USA, Nov. 28 – Dec. 2, 1994, pp. 1446–1450. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=513016

[159] ITU-R, "Guidelines for Using IMT-Advanced channel models," Geneva, Switzerland, 2009. [Online]. Available: https://www.itu.int/dms_pub/itu-r/oth/0A/06/R0A060000210001MSWE.doc

[160] MiWaveS, "Deliverable D6.1 – Detailed specification of demonstrator setups and functionality," *Unpublished*, 2014.

[161] National Instruments, "PXI platform." [Online]. Available: https://ni.com/pxi

[162] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach*, 6th ed. New York, NY, USA: Pearson Education, Inc., 2012. [Online]. Available: http://www.pearsonhighered.com/educator/product/Computer-Networking-A-TopDown-Approach-6E/9780132856201.page

[163] W. Stallings, *Data and Computer Communications*, 8th ed. Upper Saddle River, NJ, USA: Pearson Education, Inc., 2007. [Online]. Available: http://williamstallings.com/DataComm/

[164] B. A. Forouzan, *Data Communications and Networking*, 4th ed. New York, NY, USA: McGraw-Hill, 2007. [Online]. Available: http://highered.mcgraw-hill.com/sites/0072967757

[165] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits: A Design Perspective*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2003. [Online]. Available: http://icbook.eecs.berkeley.edu/

[166] W. Stallings, *Operating Systems: Internals and Design Principles*, 7th ed. Upper Saddle River, NJ, USA: Pearson Education, Inc., 2012. [Online]. Available: http://williamstallings.com/OperatingSystems/

[167] M. Steer, *Microwave and RF Design: A Systems Approach.* Raleigh, NC, USA: SciTech Publishing, Inc., 2010. [Online]. Available: http://www.scitechpublishing.com/steer_9781613530214.htm

[168] A. Luzzatto and G. Shirazi, *Wireless Transceiver Design: Mastering the Design of Modern Wireless Equipment and Systems.* Chichester, UK: John Wiley & Sons, Inc., 2007. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-047006076X.html

[169] J. L. Hennessy and D. A. Patterson, *Computer Architecture – A Quantitative Approach*, 5th ed. Waltham, MA, USA: Elsevier, 2012. [Online]. Available: http://store.elsevier.com/Computer-Architecture/John-Hennessy/isbn-9780123838728/

[170] W. Stallings, *Computer Organization and Architecture – Designing for Performance*, 9th ed. Upper Saddle River, NJ, USA: Pearson Education, Inc., 2013. [Online]. Available: http://williamstallings.com/ComputerOrganization/

[171] MiWaveS, "Deliverable D2.1 – Algorithm and functional description of heterogeneous networks," *Unpublished*, 2015.

[172] International Organization for Standardization, "ISO/IEC 7498-1: Information technology – Open Systems Interconnection – Basic reference model: The basic model," Geneva, Switzerland, 1994. [Online]. Available: http://standards.iso.org/ittf/PubliclyAvailableStandards/s020269_ISO_IEC_7498-1_1994(E).zip

[173] National Instruments, "NI PXIe-1085 – 18-slot, all-hybrid PXI Express chassis," Austin, TX, USA, 2015. [Online]. Available: http://www.ni.com/datasheet/pdf/en/ds-458

[174] ——, "NI PXIe-8135 RT – 2.3 GHz quad-core real-time embedded controller for PXI Express," Austin, TX, USA, 2014. [Online]. Available: http://www.ni.com/datasheet/pdf/en/ds-404

[175] ——, "NI 6581 – 100 MHz digital adapter module for NI FlexRIO," Austin, TX, USA, 2009. [Online]. Available: http://www.ni.com/pdf/products/us/cat_ni6581.pdf

[176] ——, "NI PXI-565x – RF and microwave signal generators with modulation capability," Austin, TX, USA, 2006. [Online]. Available: http://www.ni.com/pdf/products/us/20066815101d.pdf

[177] Active Technologies, "AT-1212 – 2-channel signal generator adapter module for NI FlexRIO," Ferrara, Italy, 2012. [Online]. Available: http://www.activetechnologies.it/000download/AT1212Specifications.pdf

[178] National Instruments, "NI 5771R – 8-bit, 3 GS/s oscilloscope adapter module for NI FlexRIO," Austin, TX, USA, 2012. [Online]. Available: http://www.ni.com/pdf/manuals/375615a.pdf

[179] ——, "NI PXIe-6674T – Timing and synchronization module," Austin, TX, USA, 2015. [Online]. Available: http://www.ni.com/datasheet/pdf/en/ds-261

[180] ——, "NI FlexRIO FPGA modules," Austin, TX, USA, 2014. [Online]. Available: http://www.ni.com/datasheet/pdf/en/ds-366

[181] ——, "NI PXIe-7975R – NI FlexRIO FPGA module for PXI Express," Austin, TX, USA, 2014. [Online]. Available: http://www.ni.com/pdf/manuals/374545a.pdf

[182] ——, "LabVIEW system design software." [Online]. Available: http://www.ni.com/labview/

[183] IEEE Instrumentation and Measurement Society, "IEEE Std 1588-2008: Standard for a precision clock synchronization protocol for networked measurement and control systems," New York, NY, USA, 2008. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4579760

[184] VIBRU, "Deliverable D2.3: Design of beamsteering algorithm and training protocol for switched-beam lens antenna," *Unpublished*, 2014.

[185] B. Yin, S. Abu-Surra, G. Xu, T. Henige, E. Pisek, Z. Pi, and J. R. Cavallaro, "High-throughput beamforming receiver for millimeter wave mobile communication," in *2013 IEEE Global Communications Conference*, Atlanta, GA, USA, Dec. 9–13, 2013, pp. 3697–3702. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6831648

[186] D. Harel and Y. Feldman, *Algorithmics – The Spirit of Computing*, 3rd ed. Harlow, UK: Addison-Wesley, 2004. [Online]. Available: http://www.springer.com/us/book/9783642272653

[187] J. Wilson, *Sensor Technology Handbook*. Amsterdam, The Netherlands: Elsevier, 2005. [Online]. Available: http://www.elsevier.com/books/sensor-technology-handbook/wilson/978-1-4933-0300-7

[188] J. Fraden, *Handbook of Modern Sensors: Physics, Designs, and Applications*, 4th ed. New York, NY, USA: Springer Publishing, 2010. [Online]. Available: http://www.springer.com/us/book/9781441964656

[189] Mathworks, "MATLAB." [Online]. Available: http://mathworks.com/products/matlab/

[190] Canonical Ltd., "Ubuntu." [Online]. Available: http://ubuntu.com/

[191] Mathworks, "randperm() – Random permutation," 2015. [Online]. Available: http://mathworks.com/help/matlab/ref/randperm.html

[192] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge, UK: Cambridge University Press, 2003. [Online]. Available: http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/probability-theory-logic-science

[193] R. Diestel, *Graph theory*. Heidelberg, Germany: Springer-Verlag, 2005. [Online]. Available: http://diestel-graph-theory.com/

[194] D. West, *Introduction to Graph Theory*, 2nd ed. New York, NY, USA: Pearson Education, Inc., 2001. [Online]. Available: https://www.pearsonhighered.com/educator/product/Introduction-to-Graph-Theory/9780130144003.page

[195] M. I. Skolnik, *Radar Handbook*, 3rd ed. New York, NY, USA: McGraw-Hill, 2008.

[196] L. N. Ridenour, *Radar System Engineering*. New York, NY, USA: McGraw-Hill, 1947.

[197] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. Cambridge, MA, USA: The MIT Press, 2009. [Online]. Available: https://mitpress.mit.edu/books/introduction-algorithms

[198] E. A. Lee and S. A. Seshia, *Introduction to Embedded Systems – A Cyber-Physical Systems Approach*, 2nd ed., 2015. [Online]. Available: http://leeseshia.org/

[199] H. D. Young, R. A. Freedman, and A. L. Ford, *University Physics with Modern Physics*, 13th ed. Boston, MA, USA: Addison-Wesley, 2010. [Online]. Available: http://www.pearsonhighered.com/educator/product/University-Physics-13E/9780321696892.page

[200] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Pearson Education, Inc., 2009. [Online]. Available: http://aima.cs.berkeley.edu/

[201] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: The MIT Press, 2012. [Online]. Available: http://mitpress-ebooks.mit.edu/product/machine-learni

[202] T. Veijalainen, "Beam steering in millimeter wave radio links for small cell mobile backhaul," M. Sc. (Tech.) thesis, Aalto University, 2014. [Online]. Available: http://urn.fi/URN:NBN:fi:aalto-201406262257

[203] B. Li, Z. Zhou, H. Zhang, and A. Nallanathan, "Efficient beamforming training for 60-GHz millimeter-wave communications: A novel numerical optimization framework," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 703–717, Feb. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6587076

[204] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer Publishing, 2006. [Online]. Available: http://www.springer.com/us/book/9780387303031

[205] O. El Ayach, R. W. Heath, and S. Rajagopal, "Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays," in *2013 IEEE Global Communications Conference*, Atlanta, GA, USA, Dec. 9–13, 2013, pp. 3476–3480. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6831611

[206] V. Venkateswaran and A.-J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5447703

[207] E. Pisek, S. Abu-Surra, J. Mott, T. Henige, and R. Sharma, "High throughput millimeter-wave MIMO beamforming system for short range communication," in *2014 IEEE Consumer Communications and Networking Conference*, Las Vegas, NV, USA, Jan. 10–13, 2014, pp. 537–543. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6866623

A Random pairing experiment

The Matlab source code used to construct Table 8 in Section 5.3 is listed below. The correct call syntax is as follows:

```
pairing(2 .^ (1:10), 2 .^ (1:10), 1000)
```

```
function varargout = pairing(varargin)
%   D = PAIRING(N, M, R) Test the number of distinct
%   ordered pairs (i, j) where 1 <= i <= N and
%   1 <= j <= M generated by aligning shuffled sequences,
%   averaging over R iterations.
%
%   Inputs:
%       N   Sequence of limits for the first sequence
%               optional, default value: 2 .^ (1:7)
%       M   Sequence of limits for the second sequence
%               optional, default value: 2 .^ (1:7)
%       R   Number of iterations
%               optional, default value: 100
%
%   Output:
%       D   Coverage structure with following fields:
%               .min = minimum
%               .max = maximum
%               .avg = number
%               .med = median
%               .std = standard deviation
%
%           Note:   Only the upper triangles of the N x M
%                   matrices is populated. All return
%                   values have been normalized by
%                   dividing them with N * M.
%
%   Author:
%       Tuomas Leinonen
%       tuomas.leinonen@aalto.fi
%       Aug 22nd, 2015
%

    % Timing
    tic;

    % Input values
    if numel(varargin) < 3
        R = uint16(100);
```

```
40          else
41              R = uint16(varargin{3});
42          end
43
44          if numel(varargin) < 2
45              M = uint16(round(2 .^ (1:7)))';
46          else
47              M = uint16(round(varargin{2}));
48          end
49
50          if numel(varargin) < 1
51              N = uint16(round(2 .^ (1:7)))';
52          else
53              N = uint16(round(varargin{1}));
54          end
55
56          % Return values
57          D = struct(...
58                  'min',      zeros(numel(N), numel(M)), ...
59                  'max',      zeros(numel(N), numel(M)), ...
60                  'avg',      zeros(numel(N), numel(M)), ...
61                  'med',      zeros(numel(N), numel(M)), ...
62                  'std',      zeros(numel(N), numel(M)) ...
63                  );
64
65          % Results per iteration
66          K = zeros(R, 1, 'double');
67
68          % Debugging information
69          T = uint32(min(numel(N), numel(M)) * ...
70              (min(numel(N), numel(M)) + 1) / 2) + ...
71              abs(numel(N) - numel(M)) * min(numel(N),
                    numel(M));
72          fprintf('Testing %d different pairs of limits %d
                times each.\n', T, R);
73          fprintf('%4s   %6s   %6s   %12s   %11s\n', ...
74              '#', 'N', 'M', 'NM', 'Time');
75          fprintf('%s\n', '-' * ones(1, 51));
76          t = uint32(0);
77
78          % Number of generation sequences
79          n = uint32(numel(N));
80          m = uint32(numel(M));
81
82          % Track tested comibations
```

```
83          W = zeros(max([n m]), min([n m]), 'uint8');
84
85          % For each sequence {1, 2, ..., N}
86          for i = uint16(1 : numel(N))
87              n = uint32(N(i));
88
89              % For each sequence {1, 2, ..., M}
90              for j = uint16(1 : numel(M))
91                  m = uint32(M(j));
92                  nm = double(double(n) * double(m));
93
94                  % Test each combination only once
95                  if W(max([i j]), min([i j])) > 0
96                      continue;
97                  end
98                  W(max([i j]), min([i j])) = 1;
99
100                 % Debugging information
101                 t = t+1;
102                 s = toc;
103                 fprintf('%4d   %6d   %6d   %12d
                        %02d:%02d:%05.2f\n', ...
104                     t, n, m, nm, ...
105                     floor(s / 3600), ...
106                     floor((s - 3600 * floor(s / 3600)) / 60),
                            ...
107                     mod(s, 60) ...
108                     );
109
110                 % Number of repeats
111                 for k = uint16(1 : R)
112                     % Create N*M pairs
113                     P = zeros(n, m, 'uint8');
114                     for u = uint32(1 : n)
115                         for v = uint32(1 : m)
116                             y = mod(m*(u-1)+v-1, n)+1;
117                             x = mod(m*(u-1)+v-1, m)+1;
118                             if y == 1
119                                 I = uint16(randperm(n))';
120                             end
121                             if x == 1
122                                 J = uint16(randperm(m))';
123                             end
124                             P(I(y), J(x)) = 1;
125                         end
```

```
126                     end
127                     % Number of distinct, i.e., unique pairs
128                     K(k) = sum(P(:));
129                 end
130 
131                 % Statistics
132                 D.min(i, j) = min(K) / nm;
133                 D.max(i, j) = max(K) / nm;
134                 D.avg(i, j) = mean(K) / nm;
135                 D.med(i, j) = median(K) / nm;
136                 D.std(i, j) = std(K) / nm;
137             end
138         end
139 
140         % Output
141         varargout = {D};
142 
143         % Timing
144         toc;
145     end
```